

```
RESULT 6
AAY49421
ID AAY49421 standard; Protein; 304 AA.
XX
AC AAY49421;
XX
DT 13-MAR-2000 (first entry)
XX
DE Human TREX1h (Exo 1) polypeptide.
XX
KW 3'-5' exonuclease; gene therapy; protocols. antineoplastic; antiviral;
KW chemotherapeutic; cell killing; human; TREX1h; Exo 1.
XX
OS Homo sapiens.
XX
PN WO9961064-A1.
XX
PD 02-DEC-1999.
XX
PF 14-MAY-1999; 99WO-US10578.
XX
PR 22-MAY-1998; 98US-0083617.
PR 12-NOV-1998; 98US-0191470.
XX
PA (UYWA-) UNIV WAKE FOREST.
XX
PI Perrino FW;
XX
DR WPI; 2000-097077/08.
DR N-PSDB; AAZ46492.
XX
PT Novel genes used in hybridization assays to detect the capacity of
PT cells to express exonucleases -
XX
PS Claim 1; Page 45-47; 93pp; English.
XX
CC The invention relates to human and mouse genes encoding 3'-5'
CC exonucleases. The exonuclease proteins of the invention can be used to
CC identify inhibitors and effectors of exonuclease activity. Specific
CC binding polypeptides, e.g. antibodies, can be used for purifying
CC exonuclease products and detection and quantification of exonuclease
CC products in fluid and tissue samples using immunological procedures.
CC Binding proteins are also useful in modulating the activity of
CC exonucleases. Polynucleotides of the invention are useful in
CC hybridization assays to detect the capacity of cells to express
CC exonucleases. They are also useful as the basis for diagnostic methods
CC useful for identifying a genetic alteration in an exonuclease locus that
CC underlies a disease state. Nucleic acids that modulate the expression of
CC the exonuclease genes, e.g. antisense nucleic acids, ribozymes, triple
CC helix oligonucleotide, can be used in gene therapy protocols. Resistance
CC or ineffectiveness of certain antineoplastic and antiviral agents may be
CC due to an exonuclease activity. A need exists for the identification of
CC metabolic factors which modulate the ability of chemotherapeutic agents
CC to effect cell killing. This need is met by the present invention. The
CC exonuclease polypeptides of the invention can be used to design and
CC identify therapeutics which increase efficiency of the chemotherapeutic
CC agent at lower doses, which are more easily tolerated in patients, and
CC reducing side effects. The present sequence represents a human TREX1h
CC (Exo 1) polypeptide.
XX
SQ Sequence 304 AA;

Query Match 80.5%; Score 33; DB 21; Length 304;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPRK 7
:|||||
Db 159 SEHGPRK 165

RESULT 7
```

```
AAG67134
ID AAG67134 standard; Protein; 369 AA.
XX
AC AAG67134;
XX
DT 13-NOV-2001 (first entry)
XX
DE Amino acid sequence of a human enzyme.
XX
KW Human; enzyme; cancer; neurological disorder; epilepsy; stroke;
KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;
KW multiple sclerosis; Parkinson's disease; amyotrophic lateral sclerosis;
KW meningitis; schizophrenia disorder; neuroskeletal disorder; allergy;
KW Addison's disease; autoimmune disease; anemia; asthma; Crohn's disease;
KW adult respiratory distress syndrome; atopic dermatitis; psoriasis;
KW diabetes mellitus; osteoporosis; pancreatitis; rheumatoid arthritis;
KW infection; genetic disorder; muscular dystrophy; Gaucher's disease;
KW Huntington's chorea; sickle cell anemia; thalassemia; atherosclerosis;
KW Von Willebrand's disease; Wilms' tumour; cell proliferative disorder;
KW leukemia; hepatitis; cirrhosis; arteriosclerosis; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 200 /note= "potential phosphorylation site"
FT Modified-site 222 /note= "potential phosphorylation site"
FT Modified-site 224 /note= "potential phosphorylation site"
FT Modified-site 245 /note= "potential phosphorylation site"
FT Modified-site 283 /note= "potential phosphorylation site"
FT Modified-site 293 /note= "potential phosphorylation site"
FT Modified-site 311 /note= "potential phosphorylation site"
FT Modified-site 318 /note= "potential phosphorylation site"
FT Modified-site 366 /note= "potential phosphorylation site"
XX WO200164896-A2.
XX 07-SEP-2001.
XX 01-MAR-2001; 2001WO-US06806.
XX 01-MAR-2000; 2000US-0186307.
XX 28-MAR-2000; 2000US-0192532.
XX 30-MAR-2000; 2000US-0193578.
XX (INCY-) INCYTE GENOMICS INC.
XX Tang YT, Lu DAM, Bandman O, Yue H, Azimzai Y, Lal P, Burford N;
XX Baughn MR;
XX WPI; 2001-550184/61.
XX N-PSDB; AAH75162.
XX Novel human enzyme molecule useful for treating and preventing, e.g.,
XX cancer, genetic disorders, neurological disorders, autoimmune and
XX inflammatory disorders -
XX Claim 1; Page 123-124; 154pp; English.
XX The present sequence represents a human enzyme. The enzyme polynucleotide
XX and polypeptide are useful for diagnosis, treatment and prevention of
XX cancers, neurological disorders (e.g. epilepsy, stroke, Alzheimer's
XX disease, Pick's disease, Huntington's disease, dementia, multiple
XX sclerosis, Parkinson's disease, amyotrophic lateral sclerosis, bacterial
XX and viral meningitis, schizophrenic disorders and neuroskeletal
```

disorders), autoimmune/inflammatory disorders (e.g. allergies, addison's disease, autoimmune diseases, adult respiratory distress syndrome, anemia, asthma, Crohn's disease, atopic dermatitis, diabetes mellitus, osteoporosis, pancreatitis, psoriasis, rheumatoid arthritis, CC and viral, bacterial, fungal, parasitic, protozoal and helminthic infections), genetic disorder (e.g. Duchenne and Becker muscular dystrophy, Gaucher's disease, Huntington's chorea, sickle cell anemia, CC thalassemia, Von Willebrand's disease and Wilms' tumour), and cell CC proliferative disorder (e.g. atherosclerosis, leukemia, hepatitis, CC cirrhosis, and arteriosclerosis). The polynucleotide is also useful in CC somatic or germline gene therapy.

XX
SQ Sequence 369 AA;

Query Match 80.5%; Score 33; DB 22; Length 369;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPRK 7
:|||||
Db 224 SEHGPRK 230

RESULT 8
ABU52987
ID ABU52987 standard; Protein; 373 AA.
XX
AC ABU52987;
XX
DT 14-APR-2003 (first entry)
XX
DE Human nucleic acid management-associated DKFZphtes3_15j3 homologue #1.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB01496.
XX
PR 18-AUG-1999; 99US-0149499.
PR 28-SEP-1999; 99US-0156503.
XX
PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies -
XX
PS Example III; Page 602-603; 1095pp; English.

XX
CC This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a homologue of a polypeptide described in the disclosure of the invention.

XX
SQ Sequence 373 AA;

Query Match 80.5%; Score 33; DB 22; Length 373;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KHGPRK 7
|||||
Db 292 KHGPKK 297

RESULT 9
ABG15782
ID ABG15782 standard; Protein; 526 AA.
XX
AC ABG15782;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #15773.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS79969.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -
XX
PS Claim 20; SEQ ID No 46141; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

XX
CC Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 526 AA;

Query Match 80.5%; Score 33; DB 22; Length 526;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TKHGPRK 7

Db 296 TKHTRK 302

RESULT 10
ABG17605

ID ABG17605 standard; Protein; 526 AA.

XX
AC ABG17605;

XX
DT 18-FEB-2002 (first entry)

XX
DE Novel human diagnostic protein #17596.

XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX
OS Homo sapiens.

XX
PN WO200175067-A2.

XX
PD 11-OCT-2001.

XX
PF 30-MAR-2001; 2001WO-US08631.

XX
PR 31-MAR-2000; 2000US-0540217.

XX
PR 23-AUG-2000; 2000US-0649167.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Drmanac RT, Liu C, Tang YT;

XX
DR WPI; 2001-639362/73.

XX
DR N-PSDB; AAS81792.

XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX
PS Claim 20; SEQ ID No 47964; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 526 AA;

Query Match 80.5%; Score 33; DB 22; Length 526;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TKHGPRK 7
| | | | |
Db 296 TKHTRK 302

RESULT 11
ABJ18912

ID ABJ18912 standard; Protein; 717 AA.

XX
AC ABJ18912;

XX
DT 06-MAR-2003 (first entry)

XX
DE Pathogen specific antigen related staphylococcal protein SEQ ID No 58.

XX
KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis.

XX
OS Staphylococcus sp.

XX
PN WO200259148-A2.

XX
PD 01-AUG-2002.

XX
PF 21-JAN-2002; 2002WO-EP00546.

XX
PR 26-JAN-2001; 2001AT-0000130.

XX
PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

XX
PI Meinke A, Nagy E, Von Ahlsen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M;
PI Tempelmaier B;

XX
DR WPI; 2003-075410/07.

XX
PT Identifying, isolating and producing hyperimmune serum-reactive
PT antigens from a pathogen, for preparing vaccine or medicament for
PT treating or preventing e.g. staphylococcal infections, comprises
PT providing antibody preparation

XX
PS Claim 24; Page 155; 252pp; English.

XX
CC The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens
CC are used in a vaccine, comprises providing antibody preparation from a
CC plasma pool of a type of animal, or individual sera with antibodies
CC against the specific pathogen, tumour, allergen, tissue or host prone to
CC auto-immunity. The hyperimmune serum-reactive antigens comprising any of
CC the 62 sequences of 53-2261 amino acids fully defined in the
CC specification, or their hyperimmune fragments are useful for the
CC manufacture of a pharmaceutical preparation, particularly a vaccine
CC against staphylococcal infections or colonisation against S. aureus or S.
CC epidermidis. The preparation of antibodies is useful for the manufacture
CC of a medicament for treating or preventing staphylococcal infections or
CC colonisation against S. aureus or S. epidermidis. The antibody
CC preparations may also be used for diagnostic and imaging purposes. Other
CC conditions that can be treated include cancer, autoimmune diseases or
CC infections caused by viral (e.g. HIV, hepatitis A, B or C), fungal or
CC protozoan pathogens. This sequence represents a staphylococcal protein
CC relating to the method for identifying and producing pathogen specific
CC antigens of the invention.

XX
SQ Sequence 717 AA;

Query Match 80.5%; Score 33; DB 24; Length 717;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKHGPR 6
| | | | |
Db 91 TKHGPK 96

RESULT 12
ABU52986
ID ABU52986 standard; Protein; 743 AA.
XX AC
XX ABU52986;
DT 14-APR-2003 (first entry)
XX DE Human nucleic acid management-associated protein from DKFZphptes3_15j3.
XX DE Human; gene therapy; vaccine; disease treatment; detection.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-IB01496.
XX PR 18-AUG-1999; 99US-0149499.
XX PR 28-SEP-1999; 99US-0156503.
XX PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX PI Wiemann S;
XX DR WPI; 2001-327840/34.
XX DR N-PSDB; ABX71347.
XX PT Nucleic acids having the sequences of clones isolated from libraries of
XX PT different human tissues, useful in recombinant DNA methodologies -
XX PS Claim 21; Page 602; 1095pp; English.
XX CC This invention describes novel polynucleotides and polypeptides isolated
XX CC from human cDNA libraries which can be used for gene therapy or in
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. The products of the
XX CC invention may also be used to identify modulators of expression and
XX CC activity and to down regulate expression and activity. The antibodies of
XX CC the invention may also be used as diagnostic agents for detecting the
XX CC presence of polypeptides in samples. This sequence represents a
XX CC polypeptide described in the disclosure of the invention.
SQ Sequence 743 AA;
Query Match 80.5%; Score 33; DB 22; Length 743;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KHGPRK 7
Db 381 KHGPKK 386
RESULT 13
AAG80778
ID AAG80778 standard; Protein; 774 AA.
XX AC
XX AAG80778;
DT 19-APR-2002 (first entry)
XX DE Human exonuclease factor NEF protein.
XX DE Human; exonuclease factor; NEF; testicular; immunisation;
XX KW antibody preparation.
XX OS Homo sapiens.
XX DR

PN CN1316433-A.
XX 10-OCT-2001.
XX 11-APR-2001; 2001CN-0108200.
XX 11-APR-2001; 2001CN-0108200.
XX (UYNA-) UNIV NANJING MEDICAL.
XX Sha J, Zhou Z, Li J;
XX WPI; 2002-067605/10.
XX DR N-PSDB; ABA97189.
XX PT Human exonuclease factor gene encoding a protein -
XX PS Claim 1; Page 2 (Disclosure); 6pp; Chinese.
XX CC This invention describes a novel human testicular exonuclease factor gene
XX CC (NEF). The NEF gene described in the disclosure if the invention can be
XX CC used to prepare a fusion protein which can be used to immunise animals to
XX CC prepare monoclonal and polyclonal antibodies. The protein of the
XX CC invention can be used to prepare medicines. This sequence represents the
XX CC human NEF protein described in the invention.
XX SQ Sequence 774 AA;
Query Match 80.5%; Score 33; DB 23; Length 774;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 KHGPRK 7
Db 381 KHGPKK 386
RESULT 14
AAU37176
ID AAU37176 standard; Protein; 897 AA.
XX AC
XX AAU37176;
DT 14-FEB-2002 (first entry)
XX DE Staphylococcus aureus cellular proliferation protein #1346.
XX KW Antisense; prokaryotic cellular proliferation protein;
XX KW antibiotic; antibacterial; drug design.
XX OS Staphylococcus aureus.
XX PN WO200170955-A2.
XX PD 27-SEP-2001.
XX PF 21-MAR-2001; 2001WO-US09180.
XX PR 21-MAR-2000; 2000US-191078P.
XX PR 23-MAY-2000; 2000US-206848P.
XX PR 26-MAY-2000; 2000US-207272P.
XX PR 23-OCT-2000; 2000US-242578P.
XX PR 27-NOV-2000; 2000US-253625P.
XX PR 22-DEC-2000; 2000US-257931P.
XX PR 16-FEB-2001; 2001US-269308P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX PI Yamamoto RT, Xu HH;
XX DR WPI; 2001-611495/70.
XX DR N-PSDB; AAS55035.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Example 3; Seq ID No 12769; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1113 AA;
Query Match 80.5%; Score 33; DB 22; Length 1113;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKHGPR 6
| | | | |
Db 523 TKHGPK 528
Search completed: February 11, 2004, 17:02:50
Job time : 34.25 secs

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX
PS Example 3; Seq ID No 12769; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 897 AA;
Query Match 80.5%; Score 33; DB 22; Length 897;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKHGPR 6
| | | | |
Db 307 TKHGPK 312
RESULT 15
AAU34340
ID AAU34340 standard; Protein; 1113 AA.
XX
AC AAU34340;
XX
DT 14-FEB-2002 (first entry)
XX
DE Staphylococcus aureus cellular proliferation protein #616.
XX
KW Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
OS Staphylococcus aureus.
XX
PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR N-PSDB; AAS52199.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	116	3	013550 saccharomyc
2	34	100.0	168	11	Q8BQA6 mus musculus
3	34	100.0	186	12	Q8JJU6 turkey rhin
4	34	100.0	186	12	Q91SF9 avian pneum
5	32	94.1	339	16	Q31702 bacillus su
6	32	94.1	478	16	Q9CJ31 lactococcus
7	32	94.1	905	10	Q8RYR1 oryza sativ
8	32	94.1	1381	16	Q8YWH7 anabaena sp
9	31	91.2	184	12	Q91R55 avian pneum
10	31	91.2	217	12	Q9YFQ9 peanut stun
11	31	91.2	217	12	Q9YPR0 peanut stun
12	31	91.2	263	16	Q8DSN0 streptococc
13	31	91.2	348	16	Q9CP10 pasteurrella
14	31	91.2	509	17	Q96XL2 sulfolobus
15	31	91.2	669	16	Q66983 aquifex aeo
16	31	91.2	780	5	Q81122 plasmodium

17	31	91.2	822	11	Q8BYW3	Q8byw3 mus musculu
18	31	91.2	2999	11	Q8CHI7	Q8chi7 mus musculu
19	31	91.2	3035	11	Q8CHI8	Q8chi8 mus musculu
20	30	88.2	151	16	Q9KAF5	Q9kaf5 bacillus ha
21	30	88.2	153	16	Q8EQL6	Q8eq16 oceanobacil
22	30	88.2	161	5	Q8WQ09	Q8wq09 ostertagia
23	30	88.2	202	8	Q9TIP1	Q9tip1 pedicularis
24	30	88.2	247	2	O30393	O30393 neisseria g
25	30	88.2	249	16	Q8EXP5	Q8exp5 leptospira
26	30	88.2	255	5	Q8WQ10	Q8wq10 ostertagia
27	30	88.2	332	17	Q9UX18	Q9ux18 sulfolobus
28	30	88.2	396	16	Q926X4	Q926x4 listeria in
29	30	88.2	396	16	Q8Y3Q6	Q8y3q6 listeria mo
30	30	88.2	447	2	Q8GDN0	Q8gdno photorhabdu
31	30	88.2	453	16	Q8D319	Q8d319 wiggleswort
32	30	88.2	620	5	Q8MXG4	Q8mxg4 caenorhabdi
33	30	88.2	723	16	Q8PH66	Q8ph66 xanthomonas
34	30	88.2	723	16	Q8P5T8	Q8p5t8 xanthomonas
35	30	88.2	735	16	Q9PGF0	Q9pgf0 xylella fas
36	30	88.2	809	10	Q8VYCI	Q8vycl arabidopsis
37	30	88.2	836	10	O24527	O24527 arabidopsis
38	30	88.2	966	5	Q8T9K4	Q8t9k4 drosophila
39	30	88.2	966	5	Q9VLM8	Q9vlm8 drosophila
40	30	88.2	966	5	Q9U6B4	Q9u6b4 drosophila
41	30	88.2	1120	10	Q9LQA1	Q9lqa1 arabidopsis
42	30	88.2	1222	10	Q8H6V1	Q8h6v1 zea mays (m
43	30	88.2	1222	10	Q8H6U6	Q8h6u6 zea mays (m
44	30	88.2	1226	10	Q8H6V0	Q8h6v0 zea mays (m
45	30	88.2	1247	10	Q8H6U9	Q8h6u9 zea mays (m

ALIGNMENTS

RESULT 1
O13550
ID O13550 PRELIMINARY; PRT; 116 AA.
AC O13550;
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE YLR269CP.
GN YLR269C.

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97313267; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansorge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Dusterhoft A.,
RA Entian K.D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Kotter P.,
RA Louis E.J., Messenguy F., Mewes H.W., Miosga T., Mostl D.,
RA Muller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portetelle D., Purnelle B., Rechmann S., Rieger M., Rinke M., Rose M.,
RA Scharfe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestarazu L.A., Vandenbol M., Verhasselt P.,
RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";
RL Nature 387:0-0(0).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Miller N.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Waterston R.;

RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Cherry J.M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U17244; AAB67386.1; -.
DR SGD; S0004259; YLR269C.
SQ SEQUENCE 116 AA; 13718 MW; 94F7025D903552E8 CRC64;

Query Match 100.0%; Score 34; DB 3; Length 116;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 67 SLKRLPK 73

RESULT 2
Q8BQA6 PRELIMINARY; PRT; 168 AA.
ID Q8BQA6
AC Q8BQA6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK051146; BAC34536.1; -.
KW Hypothetical protein.
SQ SEQUENCE 168 AA; 18003 MW; AF9BB01F4359A158 CRC64;

Query Match 100.0%; Score 34; DB 11; Length 168;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 120 SLKRLPK 126

RESULT 3
Q8JJU6 PRELIMINARY; PRT; 186 AA.
ID Q8JJU6
AC Q8JJU6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Matrix protein M2.
GN M2.
OS Turkey rhinotracheitis virus (TRTV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=11264;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UK /11 /94;
RA Jacobs J.A., Njenga K., Mawditt K., Britton P., Cavanagh D., Seal B.;
RT "Nucleotide and predicted amino acid sequence analysis of the
RT phosphoprotein, second matrix and small hydrophobic protein genes for
RT avian metapneumovirus type B.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ492378; CAD37362.1; -.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
SQ SEQUENCE 186 AA; 20873 MW; 28D549FCA5ED605D CRC64;

Query Match 100.0%; Score 34; DB 12; Length 186;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 146 SLKRLPK 152

RESULT 4
Q91SF9 PRELIMINARY; PRT; 186 AA.
ID Q91SF9
AC Q91SF9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Matrix glycoprotein M2.
GN M2.
OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=38525;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hungary/657/4;
RA Seal B.S., Jacobs J.A., Turpin E.A., Njenga M.K.;
RT "Sequence comparison of avian pneumovirus isolates from the United
RT States confirms a subgroup different strains."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF356650; AAK48888.1; -.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
SQ SEQUENCE 186 AA; 20802 MW; 6214C6187B7711B8 CRC64;

Query Match 100.0%; Score 34; DB 12; Length 186;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 146 SLKRLPK 152

RESULT 5
O31702 PRELIMINARY; PRT; 339 AA.
ID O31702
AC O31702;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Molybdopterin biosynthesis protein.
GN MOEB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

```

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=1168;
RX MEDLINE=90368558; PubMed=1697575;
RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.;
RT "Secretory S complex of Bacillus subtilis: sequence analysis and
RT identity to pyruvate dehydrogenase."
RL J. Bacteriol. 172:5052-5063(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=1168;
RX MEDLINE=97144523; PubMed=8990290;
RA Henriques A.O., Bryan E.M., Beall B.W., Moran C.P. Jr.;
RT "cse15, cse60, and csk22 are new members of mother-cell-specific
RT sporulation regulons in Bacillus subtilis."
RL J. Bacteriol. 179:389-398(1997).
DR EMBL; Z99111; CAB13300.1; -.
DR EMBL; AF012285; AAC24901.1; -.
DR InterPro; IPR000205; NAD binding.
DR InterPro; IPR000594; ThiF domain.
DR Pfam; PF00899; ThiF; 1.
KW Complete proteome.
SQ SEQUENCE 339 AA; 37517 MW; 3CEF58307DC5C343 CRC64;

Query Match 94.1%; Score 32; DB 16; Length 339;
Best Local Similarity 85.7%; Pred. No. 57;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 275 SLKRIPK 281

RESULT 6
Q9CJ31
ID Q9CJ31 PRELIMINARY; PRT; 478 AA.
AC Q9CJ31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Beta-glucosidase A (EC 3.2.1.21).
GN BGLS OR LL0175.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AF006255; AAK04273.1; -.
DR HSSP; P11546; 1PBG.
DR InterPro; IPR001360; Glyco_hydro_1.
DR Pfam; PF00232; Glyco_hydro_1; 1.
DR PRINTS; PR00131; GLHYDRLASE1.
DR ProDom; PD000650; Glyco_hydro_1; 1.
DR PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
KW Complete proteome.
SQ SEQUENCE 478 AA; 54720 MW; BE3EC9C08B0BEB1 CRC64;

Query Match 94.1%; Score 32; DB 16; Length 478;
Best Local Similarity 85.7%; Pred. No. 80;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 455 SLKRIPK 461

RESULT 7
Q8RYR1
ID Q8RYR1 PRELIMINARY; PRT; 905 AA.
AC Q8RYR1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative LRR.
GN OSJNBA0026J14.23.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC
RT clone:OSJNBA0026J14."
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP004231; BAB89236.1; -.
DR Gramene; Q8RYR1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF00931; NB-ARC; 1.
SQ SEQUENCE 905 AA; 103230 MW; F205F3B63C7484C0 CRC64;

Query Match 94.1%; Score 32; DB 10; Length 905;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 574 SLKRLPK 580

RESULT 8
Q8YWH7
ID Q8YWH7 PRELIMINARY; PRT; 1381 AA.
AC Q8YWH7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

```

DE Regulatory protein.
GN ALL1636.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003586; BAB78002.1; -
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR002182; NB-ARC.
DR Pfam; PF00931; NB-ARC; 1.
DR PRINTS; PR00364; DISEASERSIST.
KW Complete proteome.
SQ SEQUENCE 1381 AA; 152325 MW; 7EB862AFB5C33F93 CRC64;

Query Match 94.1%; Score 32; DB 16; Length 1381;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 264 SLKRLPK 270

RESULT 9
Q91R55 PRELIMINARY; PRT; 184 AA.
ID Q91R55;
AC Q91R55;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Matrix 2.
GN M2.
OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=38525;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MN/10;
RA Shin H.J., Nagaraja K.V., Halvorson D.A., Njenga M.K.;
RT "Molecular epidemiology of subgroup C avian pneumoviruses isolated
from the United States and comparison with subgroup A and B viruses.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY028554; AAK38450.1; -
DR InterPro; IPR000571; Znf CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
DR SMART; SM00356; Znf C3H1; 1.
SQ SEQUENCE 184 AA; 20864 MW; B2C0103E05011DEF CRC64;

Query Match 91.2%; Score 31; DB 12; Length 184;
Best Local Similarity 85.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 146 SLKRLPK 152

RESULT 10
Q9YPO9 PRELIMINARY; PRT; 217 AA.
ID Q9YPO9
AC Q9YPO9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Coat protein.
OS Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mi;
RA Xu Z., Higgins C.M., Chen K., Dietzgen R.G., Zhang Z., Yang L.,
RA Fang X.;
RT "Characterization of peanut stunt virus strains from China indicates a
third distinct taxonomic subgroup.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222804; CAA10998.1; -
DR InterPro; IPR000247; Cucumovirus_coat.
DR Pfam; PF00760; Cucumo_coat; 1.
DR PRINTS; PR00222; CUCUMOCOAT.
DR ProDom; PD001284; Cucumovirus_coat; 1.
SQ SEQUENCE 217 AA; 23688 MW; 573409BC46AEFDB4 CRC64;

Query Match 91.2%; Score 31; DB 12; Length 217;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 123 SLRRLPK 129

RESULT 11
Q9YPRO PRELIMINARY; PRT; 217 AA.
ID Q9YPRO
AC Q9YPRO;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Coat protein.
OS Peanut stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S;
RA Xu Z., Higgins C.M., Chen K., Dietzgen R.G., Zhang Z., Yang L.,
RA Fang X.;
RT "Characterization of peanut stunt virus strains from China indicates a
third distinct taxonomic subgroup.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222803; CAA10997.1; -
DR InterPro; IPR000247; Cucumovirus_coat.
DR Pfam; PF00760; Cucumo_coat; 1.
DR PRINTS; PR00222; CUCUMOCOAT.
DR ProDom; PD001284; Cucumovirus_coat; 1.
SQ SEQUENCE 217 AA; 23716 MW; 5734110B1C92BF3F CRC64;

Query Match 91.2%; Score 31; DB 12; Length 217;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 123 SLRRLPK 129

RESULT 12
Q8DSNO PRELIMINARY; PRT; 263 AA.
ID Q8DSNO
AC Q8DSNO;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

Job time : 26.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	186	1 VMA2 TRTV	P33494 turkey rhin
2	31	91.2	308	1 YC55 CYACA	O19888 cyanidium c
3	30	88.2	142	1 YNEK_BACSU	P45711 bacillus su
4	30	88.2	145	1 RS15_XIPMA	P70066 xiphophorus
5	30	88.2	453	1 TRME_WIGBR	Q8d319 wiggleswort
6	30	88.2	1332	1 XKDO_BACSU	P54334 bacillus su
7	29	85.3	279	1 ATPG_MYCPN	Q50330 mycoplasma
8	29	85.3	294	1 RS2_MYCPN	P75560 mycoplasma
9	29	85.3	313	1 PLDB_HAEIN	P44800 haemophilus
10	29	85.3	379	1 CYR6_MOUSE	P18406 mus musculu
11	29	85.3	379	1 CYR6_RAT	Q9es72 rattus norv
12	29	85.3	381	1 CYR6_HUMAN	O00622 homo sapien
13	29	85.3	382	1 FATB_UMBICA	Q41635 umbellulari
14	29	85.3	395	1 KIME_RAT	P17256 rattus norv
15	29	85.3	515	1 CXAA_HUMAN	P57773 homo sapien
16	29	85.3	1536	1 SIN3_YEAST	P22579 saccharomyc
17	29	85.3	2164	1 CCAA_MOUSE	P97445 mus musculu
18	29	85.3	2212	1 CCAA_RAT	P54282 rattus norv
19	29	85.3	2222	1 CCAE_RAT	Q07652 rattus norv
20	29	85.3	2223	1 CCAE_DISOM	P56699 discopoge o
21	29	85.3	2259	1 CCAE_RABIT	Q02343 cryctolagus
22	29	85.3	2272	1 CCAE_MOUSE	Q61290 mus musculu
23	29	85.3	2312	1 CCAE_HUMAN	Q15878 homo sapien
24	29	85.3	2326	1 CCAE_DISOM	P56698 discopoge o
25	29	85.3	2327	1 CCAB_MOUSE	O55017 mus musculu
26	29	85.3	2336	1 CCAB_RAT	Q02294 rattus norv
27	29	85.3	2339	1 CCAB_HUMAN	Q00975 homo sapien
28	29	85.3	2339	1 CCAB_RABIT	Q05152 oryctolagus
29	29	85.3	2424	1 CCAA_RABIT	P27884 oryctolagus
30	29	85.3	2505	1 CCAA_HUMAN	O00555 homo sapien
31	28	82.4	102	1 YW02_XANAC	Q8php7 xanthomonas
32	28	82.4	146	1 GSPG_VIBCH	P45773 vibrio chol
33	28	82.4	150	1 RL15_RICPR	Q9zcs4 rickettsia

34	28	82.4	174	1 VNSC_PHODV	P35940 phocine dis
35	28	82.4	214	1 FGP8_CHICK	Q90722 gallus gall
36	28	82.4	359	1 ADD2_STRCO	Q9x7t2 streptomyce
37	28	82.4	428	1 NH44_CABEL	Q22555 caenorhabdi
38	28	82.4	435	1 RPOD_SORBI	Q01923 sorghum bic
39	28	82.4	507	1 CBS_YEAST	P32582 saccharomyc
40	28	82.4	509	1 CHLE_MESVI	Q9mur9 mesostigma
41	28	82.4	576	1 ACH2_DROME	P17644 drosophila
42	28	82.4	686	1 RECG_STAAU	O50581 staphylococ
43	28	82.4	732	1 PRIA_ECOLI	P17888 escherichia
44	28	82.4	830	1 YBMA_SCHPO	Q10332 schizosacch
45	28	82.4	1479	1 RPOD_WHEAT	Q9xps9 triticum ae

ALIGNMENTS

RESULT 1

VMA2 TRTV

ID VMA2 TRTV STANDARD; PRT; 186 AA.

AC P33494;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Matrix glycoprotein M2 (Envelope-associated 22 kDa protein).

GN 22K.

OS Turkey rhinotracheitis virus (TRTV).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.

OX NCBI_TaxID=11264;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92333255; PubMed=1629697;

RA Ling R., Easton A.J., Pringle C.R.;

RT "Sequence analysis of the 22K, SH and G genes of turkey

rhinotracheitis virus and their intergenic regions reveals a gene

order different from that of other pneumoviruses.";

RL J. Gen. Virol. 73:1709-1715(1992).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=UK/3BV/85;

RX MEDLINE=92300329; PubMed=1607858;

RA Yu Q., Davis P.J., Brown T.D.K., Cavanagh D.;

RT "Sequence and in vitro expression of the M2 gene of turkey

rhinotracheitis pneumovirus.";

RL J. Gen. Virol. 73:1355-1363(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

CC EMBL; S40185; AAB22544.1; -.

DR EMBL; X63408; CAA45004.1; -.

DR PIR; JQ1623; JQ1623.

DR PIR; JQ1987; JQ1987.

DR InterPro; IPR000571; Znf_CCCH.

DR Pfam; PF00642; zf-CCCH; 1.

DR SMART; SM00356; Znf_C3H1; 1.

KW Matrix protein; Envelope protein; Glycoprotein.

FT CARBOHYD 86 N-LINKED (GLCNAC...) (POTENTIAL).

FT CONFLICT 175 86 N -> S (IN REF. 2).

SQ SEQUENCE 186 AA; 20986 MW; BC638B5741B612E5 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 186;

Best Local Similarity 85.7%; Pred. No. 7.4;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7

|||||:

Db 146 SLKRLPR 152

RESULT 2

YC55_CYACA STANDARD; PRT; 308 AA.

AC O19888;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical 36.5 kDa protein ycf55.

GN YCF55 OR YCF5.

OS Cyanidium caldarium.

OG Chloroplast.

OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;

OC Cyanidium.

OX NCBI_TaxID=2771;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RX-1;

RX MEDLINE=20496959; PubMed=11040290;

RA Gloeckner G., Rosenthal A., Valentin K.-U.;

RA "The structure and gene repertoire of an ancient red algal plastid genome.";

RT J. Mol. Evol. 51:382-390(2000).

RL J.

CC -!- SIMILARITY: BELONGS TO THE YCF55 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF022186; AAB82701.1; -.

DR PIR; T11956; T11956.

KW Hypothetical protein; Chloroplast.

SQ SEQUENCE 308 AA; 36498 MW; 758C719E056D859F CRC64;

Query Match 91.2%; Score 31; DB 1; Length 308;

Best Local Similarity 85.7%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 SLKRLPK 7

Db 258 SLKRLPK 264

RESULT 3

YNEK_BACSU STANDARD; PRT; 142 AA.

AC P45711;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yneK.

GN YNEK.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=97124194; PubMed=8969507;

RA Rose M., Entian K.D.;

RT "New genes in the 170 degrees region of the Bacillus subtilis genome encode DNA gyrase subunits, a thioredoxin, a xylanase and an amino acid transporter.";

RT Microbiology 142:3097-3101(1996).

RL [2]

RN

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritiz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.";

RL Nature 390:249-256(1997).

RN [3]

RP SEQUENCE OF 41-142 FROM N.A.

RC STRAIN=168;

RX MEDLINE=97221596; PubMed=9068642;

RA Schioett T., von Wachenfeldt C., Hederstedt L.;

RT "Identification and characterization of the ccdA gene, required for cytochrome c synthesis in Bacillus subtilis.";

RL J. Bacteriol. 179:1962-1973(1997).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Z73234; CAA97597.1; -.

DR EMBL; Z99113; CAB13680.1; -.

DR EMBL; X87845; CAA61119.1; -.

DR PIR; D69891; D69891.

DR Subtilist; BG11252; yneK.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 142 AA; 16989 MW; 30F697A4A4990295 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 142;

Best Local Similarity 100.0%; Pred. No. 9.4;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LKRLPK 7

Db 33 LKRLPK 38

RESULT 4

RS15_XIPMA STANDARD; PRT; 145 AA.

ID RS15_XIPMA

AC P70056;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE 40S ribosomal protein S15 (RIG protein).

GN RPS15 OR RIG.

OS Xiphophorus maculatus (Southern platyfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Cyprinodontiformes; Poeciliidae; Xiphophorus.

OX NCBI_TaxID=8083;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=JP 163A;

RA Walter R.B., Obermoeller R.D., Moore D.D., Lacson J.M.,

RA Coletta L., McEntire B.B., Morizot D.C., Nairn R.S.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL; U65408; AAB18956.1; -

HSSP; P80381; IQKF.

InterPro; IPR002222; Ribosomal_S19.

InterPro; IPR005713; S15_euk_arch.

Pfam; PF00203; Ribosomal_S19; 1.

PRINTS; PR00975; RIBOSOMALS19.

ProDom; PD001012; Ribosomal_S19; 1.

TIGRFAMS; TIGR01025; rpsS_arch; 1.

PROSITE; PS00323; RIBOSOMAL_S19; FALSE_NEG.

KW Ribosomal protein.

SQ SEQUENCE 145 AA; 16946 MW; 15896B70D21D5351 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 145;

Best Local Similarity 100.0%; Pred. No. 9.6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

DB 57 LKRLPK 62

RESULT 5

TRME_WIGBR

ID TRME_WIGBR STANDARD; PRT; 453 AA.

AC Q8D3I9;

DT 15-SEP-2003 (Rel. 42, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE tRNA modification GTPase trme.

GN TRME OR THDF OR WIGBR0120.

OS Wigglesworthia glossinidia brevipalpis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Wigglesworthia.

OX NCBI_TaxID=36870;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22297718; PubMed=12219091;

RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,

RA Aksoy S.;

RT "Genome sequence of the endocellular obligate symbiont of tsetse

RT flies, Wigglesworthia glossinidia."

RL Nat. Genet. 32:402-407(2002).

CC -!- FUNCTION: Exhibits a very high intrinsic GTPase hydrolysis rate.

CC Involved in the biosynthesis of the hypermodified nucleoside 5-

CC methylaminomethyl-2-thiouridine, which is found in the wobble

CC position of some tRNAs (By similarity).

CC -!- SIMILARITY: Belongs to the era/trme family of GTP-binding

CC proteins. Trme subfamily.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL; AB063521; BAC24158.1; -

HAMAP; MF_00379; -; 1.

InterPro; IPR005289; GTP-binding_dom.

InterPro; IPR006073; GTP1_OBG.

InterPro; IPR005225; Small_GTP.

InterPro; IPR004520; ThdF.

PRINTS; PR00326; GTP1_OBG.

TIGRFAMS; TIGR00650; MG442; 1.

TIGRFAMS; TIGR00231; small_GTP; 1.

TIGRFAMS; TIGR00450; thdF; 1.

KW tRNA processing; GTP-binding; Complete proteome.

FT NP_BIND 223 230 GTP (POTENTIAL).

FT NP_BIND 270 274 GTP (POTENTIAL).

FT NP_BIND 333 336 GTP (POTENTIAL).

SQ SEQUENCE 453 AA; 50774 MW; 58C06F24BED7A7CE CRC64;

Query Match 88.2%; Score 30; DB 1; Length 453;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

DB 37 LKRLPK 42

RESULT 6

XKDO_BACSU

ID XKDO_BACSU STANDARD; PRT; 1332 AA.

AC P54334;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phage-like element PBSX protein xkdo.

GN XKDO.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RA Krogh S., O'Reilly M., Nolan N., Devine K.M.;

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,

RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,

RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier P., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- SIMILARITY: STRONG, TO B.SUBTILIS YQBO.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z70177; CAA94037.1; -;
DR EMBL; Z99110; CAB13125.1; -;
DR PIR; F69732; F69732.
DR Subtilist; BG11549; xkdo.
DR InterPro; IPR000189; SLT_domain.
DR Pfam; PF01464; SLT; 1.
KW Complete proteome.
SQ SEQUENCE 1332 AA; 145149 MW; A96C9CF9E31DF01 CRC64;
Query Match 88.2%; Score 30; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred.No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LKRLPK 7
DB 89 LKRLPK 94
RESULT 7
ATPG_MYCPN STANDARD; PRT; 279 AA.
ID ATPG_MYCPN
AC Q50330;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase gamma chain (EC 3.6.3.14).
GN ATPG OR MPN599 OR MP243.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT cluster of ribosomal protein genes.";
RL Nucleic Acids Res. 24:628-639(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON.
CC GRADIENT ACROSS THE MEMBRANE. THE GAMMA CHAIN IS BELIEVED TO BE
CC IMPORTANT IN REGULATING ATPASE ACTIVITY AND THE FLOW OF PROTONS

CC THROUGH THE CF(0) COMPLEX.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -!- SIMILARITY: Belongs to the ATPase gamma chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U43738; AAC43658.1; -;
DR EMBL; AE000024; AAB95891.1; -;
DR PIR; S62848; S62848.
DR InterPro; IPR000131; ATPase_gamma.
DR Pfam; PF00231; ATP-synt; 1.
DR PRINTS; PR00126; ATPASEGAMMA.
DR TIGRFAMs; TIGR01146; ATPsyn_Flgamma; 1.
DR PROSITE; PS00153; ATPASE_GAMMA; FALSE_NEG.
KW ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase;
KW Complete proteome.
SQ SEQUENCE 279 AA; 32394 MW; D3093D376DE2C89E CRC64;
Query Match 85.3%; Score 29; DB 1; Length 279;
Best Local Similarity 71.4%; Pred.No. 33;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
DB 59 SLKRWPK 65
RESULT 8
RS2_MYCPN STANDARD; PRT; 294 AA.
ID RS2_MYCPN
AC P75560;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2.
GN RPSB OR MPN208 OR MP623.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000059; AAB96271.1; -;
DR PIR; S73949; S73949.
DR HAMAP; MF_00291; -; 1.
DR InterPro; IPR001865; Ribosomal_S2.

```
DR InterPro; IPR005706; S2 bact_org.
DR Pfam; PF00318; Ribosomal_S2; 1.
DR PRINTS; PR00395; RIBOSOMALS2.
DR TIGRFAMs; TIGR01011; rpsB_bact; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; FALSE_NEG.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 294 AA; 33425 MW; 60D111400C5F350D CRC64;

Query Match      85.3%; Score 29; DB 1; Length 294;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 6
Db      168 SLKRLPK 173

RESULT 9
PLDB_HAEIN STANDARD; PRT; 313 AA.
AC P44800;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable lysophospholipase L2 (EC 3.1.1.5) (Lecithinase B).
GN PLDB OR HI0645.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."
RL Science 269:496-512(1995).
CC -!- CATALYTIC ACTIVITY: 2-lysophosphatidylcholine + H(2)O =
CC glycerophosphocholine + a fatty acid anion.
CC -!- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
Db      237 AIKRLPK 243

::|||||
RESULT 10
CYR6_MOUSE
ID CYR6_MOUSE STANDARD; PRT; 379 AA.
AC P18406;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10) (3CH61).
GN CYR61 OR IGFBP10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Fibroblast;
RX MEDLINE=90287146; PubMed=2355916;
RA O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
RT "Expression of cyr61, a growth factor-inducible immediate-early
RT gene."
RL Mol. Cell. Biol. 10:3569-3577(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AJ; TISSUE=Embryonic fibroblast;
RX MEDLINE=91288203; PubMed=2062642;
RA Latinkic B.V., O'Brien T.P., Lau L.F.;
RT "Promoter function and structure of the growth factor-inducible
RT immediate early gene cyr61."
RL Nucleic Acids Res. 19:3261-3267(1991).
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
CC IN LUNG.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
CC -!- INDUCTION: By growth factors.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```


DR SMART; SM00214; VWC; 1. 24
DR PROSITE; PS01185; CTCK 1; 1. 164
DR PROSITE; PS01225; CTCK_2; 1. 271
DR PROSITE; PS00222; IGF BINDING; 1. 358
DR PROSITE; PS00092; TSP1; 1. 321
DR PROSITE; PS01208; VWFC 1; 1. 335
DR PROSITE; PS0184; VWFC_2; 1. 351
DR PROSITE; PS0184; VWFC_2; 1. 353
KW Growth factor binding; Signal.
FT SIGNAL 1 24
FT CHAIN 25 379
FT DOMAIN 98 164
FT DOMAIN 226 271
FT DOMAIN 284 358
FT DISULFID 284 321
FT DISULFID 301 335
FT DISULFID 312 351
FT DISULFID 315 353
FT DISULFID 320 357
SQ SEQUENCE 379 AA; 41709 MW; FA6B5014B56A8EE9 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
Db 202 SLKRLP 207

RESULT 11
CYR6 RAT
ID -CYR6 RAT STANDARD; PRT; 379 AA.
AC Q9ES72;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10).
GN CYR61 OR IGFBP10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=20435857; PubMed=10852911;
RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M., Nitsch R.M.;
RA "Muscarinic acetylcholine receptors induce the expression of the immediate early growth regulatory gene CYR61.";
RL J. Biol. Chem. 275:28929-28936(2000).
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF218568; AAC14964.1; -.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insi_gro_fac_pr.

DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWFC_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK 1; 1. 164
DR PROSITE; PS01225; CTCK_2; 1. 271
DR PROSITE; PS00222; IGF BINDING; 1. 358
DR PROSITE; PS00092; TSP1; 1. 321
DR PROSITE; PS01208; VWFC 1; 1. 335
DR PROSITE; PS0184; VWFC_2; 1. 351
DR PROSITE; PS0184; VWFC_2; 1. 353
KW Growth factor binding; Signal.
FT SIGNAL 1 24
FT CHAIN 25 379
FT DOMAIN 98 164
FT DOMAIN 226 271
FT DOMAIN 284 358
FT DISULFID 284 321
FT DISULFID 301 335
FT DISULFID 312 351
FT DISULFID 315 353
FT DISULFID 320 357
SQ SEQUENCE 379 AA; 41687 MW; 62BFOBBA4C5AFDE9 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
Db 202 SLKRLP 207

RESULT 12
CYR5 HUMAN
ID -CYR5 HUMAN STANDARD; PRT; 381 AA.
AC O00622; O14934; O43775; Q9BZL7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10) (GIG1 protein).
GN CYR61 OR IGFBP10 OR GIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97280750; PubMed=9135077;
RA Jay P., Berge-lefranc J.L., Marsollier C., Mejean C., Taviaux S., Berta P.;
RA "The human growth factor-inducible immediate early gene, CYR61, maps to chromosome 1p.";
RL Oncogene 14:1753-1757(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98197344; PubMed=9536281;
RA Martinerie C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
RT "Chromosomal mapping and expression of the human cyr61 gene in tumour cells from the nervous system.";
RL Mol. Pathol. 50:310-316(1997).
RN [4]

RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kolesnikova T.V., Lau L.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Bi A.B., Yu L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Schuetze N., Lechner A., Groll C., Koehrle J., Jakob F.;
RT "Regulation of hCYR61 by vitamin D, serum and cytokines in fetal human
osteoblasts";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
RT "Organization and expression of the CYR61 gene in normal human
fibroblasts";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung, Placenta, and Skin;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
CC PROTEINS: PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Contains 1 IGFBP domain.
CC -!- SIMILARITY: Contains 1 WFPC domain.
CC -!- SIMILARITY: Contains 1 TSP type-1 domain.
CC -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y12084; CAA72802.1; -.
DR EMBL; U62015; AAB58319.1; -.
DR EMBL; Y11307; CAA72167.1; -.
DR EMBL; AF003594; AAB61240.1; -.
DR EMBL; AF031385; AAB84227.1; -.
DR EMBL; Z98053; CAB10848.1; -.
DR EMBL; AF307860; AAG59863.1; -.
DR EMBL; BC001271; AAH01271.1; -.
DR EMBL; BC009199; AAH09199.1; -.
DR EMBL; BC016952; AAH16952.1; -.
DR Genew; HGNC:2654; CYR61.
DR MIM; 602369; -.

DR GO; GO:0008283; P:cell proliferation; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; VWF_1; 1.
DR PROSITE; PS0184; VWF_2; 1.
KW Growth factor binding; Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 381 CYR61 PROTEIN.
FT DOMAIN 26 97 IGFBP.
FT DOMAIN 98 164 VWF.
FT DOMAIN 228 273 TSP_TYBPB-1.
FT DOMAIN 286 360 CTCK.
FT DISULFID 286 323 BY SIMILARITY.
FT DISULFID 303 337 BY SIMILARITY.
FT DISULFID 314 353 BY SIMILARITY.
FT DISULFID 317 355 BY SIMILARITY.
FT DISULFID 322 359 BY SIMILARITY.
FT CONFLICT 165 165 E -> Q (IN REF. 3).
FT CONFLICT 210 210 L -> I (IN REF. 5).
FT CONFLICT 220 220 L -> R (IN REF. 5).
FT CONFLICT 369 369 F -> L (IN REF. 7).
SQ SEQUENCE 381 AA; 42026 MW; FCOBD39C078CA0B1 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 381;
Best Local Similarity 100.0%; Pred.No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
Db 206 SLKRLP 211
|||||

RESULT 13
FATE_UMBCA
ID_FATB_UMBCA STANDARD; PRT; 382 AA.
AC Q41635;
DT 30-MAY-2000 (Rel. 39; Created)
DT 30-MAY-2000 (Rel. 39; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Lauroyl-acyl carrier protein thioesterase, chloroplast precursor
DE (EC 3.1.2.14) (12:0-acyl-carrier protein thioesterase) (12:0-ACP
DE thioesterase) (Acyl-[acyl-carrier protein] hydrolase) (BTE).
GN FATB1 OR FATB.
OS Umbellularia californica (California bay).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Umbellularia.
OX NCBI_TaxID=3438;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92320297; PubMed=1621095;
RA Voelker T.A., Worrell A.C., Anderson L., Bleibaum J., Fan C.,
RA Hawkins D.J., Radke S.E., Davies H.M.;
RT "Fatty acid biosynthesis redirected to medium chains in transgenic
oilseed plants";
RL Science 257:72-74 (1992).
RN [2]
RP MUTAGENESIS TO CHANGE SPECIFICITY.

RX MEDLINE=96068671; PubMed=7479856;
RA Yuan L., Voelker T.A., Hawkins D.J.;
RT "Modification of the substrate specificity of an acyl-acyl carrier
RT protein thioesterase by protein engineering";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10639-10643(1995).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CHAIN TERMINATION DURING DE
CC NOVO FATTY ACID SYNTHESIS. HIGH THIOESTERASE ACTIVITY FOR LAUROYL-
CC ACP VS. OTHER ACYL-ACPS.
CC -!- CATALYTIC ACTIVITY: Lauroyl-[acyl-carrier protein] + H(2)O =
CC [acyl-carrier protein] + laurate.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- BIOTECHNOLOGY: Introduced by genetic manipulation and expressed in
CC canola by Monsanto (Calgene) so as to obtain laurate-rich seeds.
CC -!- SIMILARITY: BELONGS TO THE ACYL-ACP THIOESTERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M94159; AAA34215.1; --
CC PIR; A40229; A40229.
CC InterPro; IPR002864; Acyl-ACP_TE.
CC Pfam; PF01643; Acyl-ACP TE; 1.
KW Fatty acid biosynthesis; Hydrolase; Chloroplast; Transit peptide;
KW Genetically modified food.
FT TRANSIT 1 60 CHLOROPLAST (POTENTIAL).
FT CHAIN 61 382 LAUROYL-ACYL CARRIER PROTEIN
FT THIOESTERASE.
FT MUTAGEN 197 197 M->R: CONVERTED TO A 14:0 ACP TE; WHEN
FT ASSOCIATED WITH H-199 AND K-231.
FT MUTAGEN 199 199 R->H: CONVERTED TO A 14:0 ACP TE; WHEN
FT ASSOCIATED WITH R-197 AND K-231.
FT MUTAGEN 231 231 T->K: CONVERTED TO A 14:0 ACP TE; WHEN
FT ASSOCIATED WITH R-197 AND H-199.
FT SEQUENCE 382 AA; 42915 MW; 96262A558545E89F CRC64;
Query Match 85.3%; Score 29; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLP 6
Db 56 SLKRLP 61
RESULT 14
KIME RAT
ID KIME RAT STANDARD; PRT; 395 AA.
AC P17256;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mevalonate kinase (EC 2.7.1.36) (MK).
GN MVK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=90222132; PubMed=2158094;
RA Tanaka R.D., Lee L.Y., Schafer B.L., Kratunis V.J., Mohler W.A.,
RA Robinson G.W., Mosley S.T.;
RT "Molecular cloning of mevalonate kinase and regulation of its mRNA
RT levels in rat liver."
RL Proc. Natl. Acad. Sci. U.S.A. 87:2872-2876(1990).
CC -!- FUNCTION: MAY BE A REGULATORY SITE IN CHOLESTEROL BIOSYNTHETIC
CC PATHWAY.

CC -!- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-
CC phosphomevalonate.
CC -!- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE
CC COMPETITIVE INHIBITORS.
CC -!- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.
CC -!- DISEASE: MUTATION IN THE MEVALONATE KINASE GENE CAUSES MEVALONIC
CC ACIDURIA.
CC -!- SIMILARITY: BELONGS TO THE GHMP KINASE FAMILY. MEVALONATE KINASE
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M29472; AAA41588.1; --
CC PIR; A35629; A35629.
CC PDB; 1KVK; 19-JUN-02.
CC InterPro; IPR001174; Galkinase.
CC InterPro; IPR006204; GHMP_kinase.
CC InterPro; IPR006203; GHMP_knse_Atp.
CC InterPro; IPR006205; Mv_gal_Kin.
CC InterPro; IPR006206; Mv_galk_kinase.
CC Pfam; PF00288; GHMP_kinases; 1.
CC PRINTS; PR00960; LMEPPROTEIN.
CC PRINTS; PR00959; MEVGALKINASE.
CC TIGRFAms; TIGR00549; mevalon_kin; 1.
CC PROSITE; PS00627; GHMP_KINASES_ATP; 1.
KW Transferase; Kinase; Cholesterol biosynthesis; ATP-binding;
KW Peroxisome; 3D-structure.
FT NP BIND 138 148 ATP (POTENTIAL).
FT SEQUENCE 395 AA; 41987 MW; 803D1F44E3C525FC CRC64;
Query Match 85.3%; Score 29; DB 1; Length 395;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLP 6
Db 223 SLKRLP 228
RESULT 15
CXAA HUMAN
ID CXAA HUMAN STANDARD; PRT; 515 AA.
AC P57773;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gap junction alpha-10 protein (Connexin 59) (Cx59).
GN GJA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Coucke P.J., Van Laer L., Meyers J., Van Hauwe P., Ottschytsch N.,
RA Wauters G., Kelley P., Willems P.J., Van Camp G.;
RT "Identification of a new connexin gene using degenerate PCR
RT primers."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -!- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. ALPHA-TYPE (GROUP II)

```
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF179597; AAC09406.1; -.
CC Genew; HGNC:19155; GJA10.
CC InterPro; IPR000500; Connexin.
CC Pfam; PF00029; connexin; 1.
CC PRINTS; PR00206; CONNEXIN.
CC SMART; SM00037; CNX; 1.
CC PROSITE; PS00407; CONNEXINS_1; 1.
CC PROSITE; PS00408; CONNEXINS_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 20 40 POTENTIAL.
FT DOMAIN 41 77 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 78 98 POTENTIAL.
FT DOMAIN 99 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 187 POTENTIAL.
FT DOMAIN 188 209 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 210 230 POTENTIAL.
FT DOMAIN 231 515 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 515 AA; 58856 MW; EEBF0E9510AD0C96 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 515;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SLKRLP 6
Db 274 SLKRLP 279
```

Search completed: February 11, 2004, 17:04:11
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	34	100.0	116	2 S70038	hypothetical prote
2	32	94.1	339	2 F69659	molybdopterin bios
3	32	94.1	478	2 G86646	beta-glucosidase (
4	32	94.1	1381	2 AF2010	regulatory protein
5	31	91.2	186	1 JQ1623	envelope-associate
6	31	91.2	186	2 JQ1987	hypothetical 21.0K
7	31	91.2	308	2 T11956	hypothetical prote
8	31	91.2	669	2 D70369	ATP-dependent DNA
9	30	88.2	142	2 D69891	yneK protein - Bac
10	30	88.2	151	2 D83941	hypothetical prote
11	30	88.2	332	2 H90209	GTP-binding protei
12	30	88.2	396	2 AE1796	efflux protein hom
13	30	88.2	396	2 AH1421	efflux protein hom
14	30	88.2	735	2 A82817	pentaphosphate gua
15	30	88.2	836	2 B96716	probable serine/th
16	30	88.2	1332	2 F69732	PBSX prophage ORF
17	30	88.2	1590	2 B86398	protein T7N9.24 [i
18	29	85.3	173	2 I51269	lutropin receptor
19	29	85.3	273	2 G89856	conserved hypothet
20	29	85.3	279	2 S62848	H+-transporting tw
21	29	85.3	294	2 S73949	ribosomal protein
22	29	85.3	313	2 A64084	lysophospholipase
23	29	85.3	333	2 T41669	hypothetical prote
24	29	85.3	362	2 S69197	oleoyl-[acyl-carri
25	29	85.3	362	2 T14393	probable oleoyl-[a
26	29	85.3	366	2 S40408	oleoyl-[acyl-carri
27	29	85.3	366	2 S40407	oleoyl-[acyl-carri
28	29	85.3	367	2 T10207	oleoyl-[acyl-carri
29	29	85.3	379	2 A35669	gene CYR61 protein

30	29	85.3	382	2 A40229	oleoyl-lacyl-carri
31	29	85.3	395	1 A35629	mevalonate kinase
32	29	85.3	477	2 G86833	6-phospho-beta-glu
33	29	85.3	482	2 G83928	hypothetical prote
34	29	85.3	491	2 JE0396	phospho-beta-galac
35	29	85.3	522	2 T31310	hypothetical prote
36	29	85.3	537	2 D70478	conserved hypothet
37	29	85.3	589	2 HB3085	conserved hypothet
38	29	85.3	732	2 AD0014	primosomal protein
39	29	85.3	1536	1 RGBYS3	regulatory protein
40	29	85.3	2178	2 S29237	calcium channel pr
41	29	85.3	2198	2 T20371	hypothetical prote
42	29	85.3	2212	2 A41098	calcium channel pr
43	29	85.3	2222	2 A37490	voltage-dependent
44	29	85.3	2223	2 A47447	calcium channel pr
45	29	85.3	2237	2 T45115	N-type calcium cha

ALIGNMENTS

RESULT 1

S70038
hypothetical protein YLR269c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L8479.3.a
C;Species: Saccharomyces cerevisiae
C;Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 19-Apr-2002
C;Accession: S70038
R;Miller, N.
submitted to the EMBL Data Library, November 1994
A;Description: The sequence of S. cerevisiae cosmid 8479.
A;Reference number: S51395
A;Accession: S70038
A;Molecule type: DNA
A;Residues: 1-116 <MIL>
A;Cross-references: EMBL:U17244; NID:9577171; PIDN:AAB67386.1; PID:92340975; GSPDB:GN00
C;Genetics:
A;Gene: MIPS:YLR269c
A;Cross-references: SGD:S0004259
A;Map position: 12R
C;Superfamily: Saccharomyces hypothetical protein YLR269c

Query Match 100.0%; Score 34; DB 2; Length 116;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 67 SLKRLPK 73

RESULT 2

F69659
molybdopterin biosynthesis protein moeB - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C;Accession: F69659
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlor
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serc
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: F69659
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-339 <KUN>
A;Cross-references: GB:Z99111; GB:AL009126; NID:G2633699; PIDN:CAE13300.1; PID:ell185017;
A;Experimental source: strain 168
C;Genetics:
A;Gene: moeB

Query Match 94.1%; Score 32; DB 2; Length 339;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||:
Db 275 SLKRIPK 281

RESULT 3
G86646
beta-glucosidase (EC 3.2.1.21) [imported] - Lactococcus lactis subsp. lactis (strain IL1
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: G86646
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: G86646
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-478 <STO>
A;Cross-references: GB:AE005176; PID:gl2723027; PIDN:AAK04273.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: bglS
C;Superfamily: Agrobacterium beta-glucosidase
C;Keywords: glycosidase; hydrolase

Query Match 94.1%; Score 32; DB 2; Length 478;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||:
Db 455 SLKRIPK 461

RESULT 4
AF2010
regulatory protein all1636 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AF2010
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AF2010
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1381 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAE78002.1; PID:gl17135456; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1636

Query Match 94.1%; Score 32; DB 2; Length 1381;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||:
Db 264 SLKRMPK 270

RESULT 5
JQ1623
envelope-associated 22K protein - turkey rhinotracheitis virus
C;Species: turkey rhinotracheitis virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: JQ1623
R;Ling, R.; Easton, A.J.; Pringle, C.R.
J. Gen. Virol. 73, 1709-1715, 1992
A;Title: Sequence analysis of the 22K, SH and G genes of turkey rhinotracheitis virus a
A;Reference number: PQ0405; MUID:92333255; PMID:1629697
A;Accession: JQ1623
A;Molecule type: mRNA
A;Residues: 1-186 <LIN>
A;Cross-references: GB:S40185; NID:G251600; PIDN:AAE22544.1; PID:G251602
C;Genetics:
A;Gene: 22K
C;Superfamily: respiratory syncytial virus envelope-associated 22K protein
C;Keywords: glycoprotein
F;86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.2%; Score 31; DB 1; Length 186;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||:
Db 146 SLKRLLR 152

RESULT 6
JQ1987
hypothetical 21.0K protein - turkey rhinotracheitis virus (strain UK/3BV/85)
N;Alternate names: ORF 1 protein
C;Species: turkey rhinotracheitis virus
C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 28-May-1999
C;Accession: JQ1987
R;Yu, Q.; Davis, P.J.; Brown, T.D.K.; Cavanagh, D.
J. Gen. Virol. 73, 1355-1363, 1992
A;Title: Sequence and in vitro expression of the M2 gene of turkey rhinotracheitis pneu
A;Reference number: JQ1987; MUID:92300329; PMID:1607858
A;Accession: JQ1987
A;Molecule type: mRNA
A;Residues: 1-186 <YUQ>
A;Cross-references: GB:X63408; NID:G297846; PIDN:CAA45004.1; PID:G297847
C;Superfamily: respiratory syncytial virus envelope-associated 22K protein

Query Match 91.2%; Score 31; DB 2; Length 186;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||||:
Db 146 SLKRLLR 152

RESULT 7
T11956
hypothetical protein ORF308 - red alga (Cyanidium caldarium) chloroplast
C;Species: chloroplast Cyanidium caldarium
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 05-May-2000
C;Accession: T11956
R;Gloeckner, G.; Rosenthal, A.; Valentin, K.
submitted to the EMBL Data Library, September 1997
A;Description: Organisation of 46 kb of the Cyanidium caldarium RK1 plastid genome.
A;Reference number: Z17374
A;Accession: T11956
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-308 <GLO>
A;Cross-references: EMBL:AF022186; NID:g2465730; PIDN:AAB82701.1; PID:g2465772
A;Experimental source: strain RK1
C;Genetics:
A;Genome: chloroplast
A;Note: ycf5
C;Keywords: chloroplast

Query Match 91.2%; Score 31; DB 2; Length 308;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:||||
Db 258 SLKKLPK 264

RESULT 8
D70369
ATP-dependent DNA helicase REP - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C;Accession: D70369
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.

Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70369
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-669 <AQF>
A;Cross-references: GB:AE000708; NID:g29833356; PIDN:AAC06949.1; PID:g29833362; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: rep
C;Superfamily: helicase II

Query Match 91.2%; Score 31; DB 2; Length 669;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:||||
Db 417 SLKKLPK 423

RESULT 9
D69891
yneK protein - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C;Accession: D69891; S57765
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D69891
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-142 <KUN>

A;Cross-references: GB:Z99113; GB:ALJ009126; NID:g2634090; PIDN:CAB13680.1; PID:g2634180
A;Experimental source: strain 168
R;Schiott, T.; von Wachenfeldt, C.; Hederstedt, L.
submitted to the EMBL Data Library, June 1995
A;Reference number: S57401
A;Accession: S57765
A;Molecule type: DNA
A;Residues: 41-142 <SCH>
A;Cross-references: EMBL:X87845; NID:g870921; PIDN:CAA61119.1; PID:g870928
A;Experimental source: strain 168
C;Genetics:
A;Gene: yneK
C;Superfamily: Bacillus subtilis yneK protein

Query Match 88.2%; Score 30; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7
|||:||||
Db 33 LKRLPK 38

RESULT 10
D83941
hypothetical protein BH2332 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: D83941
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: D83941
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-151 <STO>
A;Cross-references: GB:AP001515; GB:BA000004; NID:gl0174886; PIDN:BAB06051.1; GSPDB:GNO
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2332
C;Superfamily: Bacillus subtilis yneK protein

Query Match 88.2%; Score 30; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7
|||:||||
Db 32 LKRLPK 37

RESULT 11
H90209
GTP-binding protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: H90209
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: H90209
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <KUR>
A;Cross-references: GB:AE006641; NID:gl3813793; PIDN:AAK40935.1; GSPDB:GN00155
C;Genetics:
A;Gene: SSO0625
C;Superfamily: GTP-binding protein, GTP1/OBG; translation elongation factor Tu homology

Query Match 88.2%; Score 30; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LKRLPK 7
Db 22 LKRLPK 27

RESULT 12
AE1796
efflux protein homolog lin2916 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AE1796
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AE1796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-396 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC98141.1; PID:gl6415457; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2916

Query Match 88.2%; Score 30; DB 2; Length 396;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 181 SLKRLPE 187

RESULT 13
AH1421
efflux protein homolog lmo2777 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1421
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-396 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00990.1; PID:gl6412277; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2777

Query Match 88.2%; Score 30; DB 2; Length 396;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 181 SLKRLPE 187

RESULT 14
A82817
pentaphosphate guanosine-3'-pyrophosphohydrolase XF0352 [imported] - Xylella fastidiosa
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C;Accession: A82817
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: A82817
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-735 <SIM>
A;Cross-references: GB:AE003887; GB:AE003849; NID:g9105175; PIDN:AAF83162.1; GSPDB:GN00
A;Experimental source: strain 9a5C
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
., F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0352
C;Superfamily: guanosine 3',5'-bis(diphosphate) 3'-pyrophosphatase

Query Match 88.2%; Score 30; DB 2; Length 735;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SLKRLPK 7
Db 532 SLERLPK 538

RESULT 15
B96716
probable serine/threonine kinase F23O10.20 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: B96716
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B96716
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-836 <STO>
A;Cross-references: GB:AE005173; NID:g7705100; PIDN:AAF67779.1; GSPDB:GN00141
C;Genetics:
A;Gene: F23O10.20
A;Map position: 1

Query Match 88.2%; Score 30; DB 2; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LKRLPK 7
| | | | |
Db 81 LKRLPK 86

Search completed: February 11, 2004, 17:11:41
Job time : 10.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-7
Perfect score: 34
Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	7	10	US-09-901-187B-7
2	32	94.1	478	12	Sequence 7, Appli
3	31	91.2	45	11	Sequence 18305, A
4	31	91.2	669	12	Sequence 3818, Ap
5	30	88.2	62	10	Sequence 58, Appl
6	30	88.2	62	10	Sequence 1778, Ap
7	30	88.2	219	16	Sequence 1778, Ap
8	29	85.3	178	12	Sequence 40, Appl
9	29	85.3	271	9	Sequence 3713, Ap
10	29	85.3	273	8	Sequence 5295, Ap
11	29	85.3	273	8	Sequence 5196, Ap
12	29	85.3	273	9	Sequence 12548, A
13	29	85.3	362	15	Sequence 20, Appl
14	29	85.3	374	9	Sequence 12, Appl
15	29	85.3	375	11	Sequence 7, Appli
	29	85.3	379	9	Sequence 11, Appl

16	29	85.3	379	12	US-10-099-322-45	Sequence 45, Appl
17	29	85.3	379	12	US-10-182-432-2	Sequence 2, Appli
18	29	85.3	379	12	US-10-044-564-45	Sequence 45, Appl
19	29	85.3	379	14	US-10-053-753-2	Sequence 2, Appli
20	29	85.3	381	11	US-09-901-910-2	Sequence 5, Appli
21	29	85.3	381	12	US-10-394-015-5	Sequence 5, Appli
22	29	85.3	381	12	US-10-099-322-42	Sequence 42, Appl
23	29	85.3	381	12	US-10-099-322-43	Sequence 43, Appl
24	29	85.3	381	12	US-10-099-322-44	Sequence 44, Appl
25	29	85.3	381	12	US-10-182-432-4	Sequence 4, Appli
26	29	85.3	381	12	US-10-044-564-42	Sequence 42, Appl
27	29	85.3	381	12	US-10-044-564-43	Sequence 43, Appl
28	29	85.3	381	12	US-10-044-564-44	Sequence 44, Appl
29	29	85.3	381	14	US-10-053-753-4	Sequence 4, Appli
30	29	85.3	381	15	US-10-294-796-2	Sequence 2, Appli
31	29	85.3	382	15	US-10-100-121-10	Sequence 10, Appl
32	29	85.3	455	9	US-09-925-301-1432	Sequence 1432, Ap
33	29	85.3	477	12	US-10-369-493-18471	Sequence 18471, A
34	29	85.3	515	12	US-09-863-776-47	Sequence 47, Appl
35	29	85.3	522	12	US-10-029-120-8	Sequence 8, Appli
36	29	85.3	522	14	US-10-027-806-8	Sequence 8, Appli
37	29	85.3	522	14	US-10-034-623-8	Sequence 8, Appli
38	29	85.3	522	15	US-10-027-801-8	Sequence 8, Appli
39	29	85.3	538	15	US-10-013-803A-2	Sequence 2, Appli
40	29	85.3	546	12	US-10-094-749-1953	Sequence 1953, Ap
41	29	85.3	582	12	US-10-108-260A-3683	Sequence 3683, Ap
42	29	85.3	642	12	US-10-369-493-12128	Sequence 12128, A
43	29	85.3	1478	12	US-10-334-143-18	Sequence 18, Appl
44	29	85.3	1536	10	US-09-801-368-322	Sequence 322, App
45	29	85.3	2198	12	US-10-369-493-5496	Sequence 5496, Ap

ALIGNMENTS

RESULT 1
US-09-901-187B-7
; Sequence 7, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatmen
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-7

Query Match 100.0%; Score 34; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
| | | | |
Db 1 SLKRLPK 7

RESULT 2
US-10-369-493-18305
; Sequence 18305, Application US/10369493

```
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18305
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-10-369-493-18305

Query Match          94.1%; Score 32; DB 12; Length 478;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      |||||
Db      455 SLKRIPK 461

RESULT 3
US-09-764-891-3818
; Sequence 3818, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3818
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3818

Query Match          91.2%; Score 31; DB 11; Length 45;
Best Local Similarity 85.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      :|||||
Db      29 ALKRLPK 35

RESULT 4
US-10-369-493-58
; Sequence 58, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 58
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Aquifex aeolicus
US-10-369-493-58

Query Match          91.2%; Score 31; DB 12; Length 669;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
      |||||
Db      417 SLKKLPK 423

RESULT 5
US-09-764-877-1778
; Sequence 1778, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1778
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-877-1778

Query Match          88.2%; Score 30; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
      |||||
Db      20 LKRLPK 25

RESULT 6
US-10-242-515-1778
; Sequence 1778, Application US/10242515
; Publication No. US20040009488A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005C1
; CURRENT APPLICATION NUMBER: US/10/242,515
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,877
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
```

```

; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1778
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-242-515-1778

Query Match      88.2%; Score 30; DB 12; Length 62;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
Db      20 LKRLPK 25

RESULT 7
US-10-176-306-40
; Sequence 40, Application US/10176306
; Publication No. US20030130485A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
; FILE REFERENCE: 10448-195001
; CURRENT APPLICATION NUMBER: US/10/176,306
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: 10/001,137
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: PCT/US01/45291
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/248,362
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,331
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/248,365
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: 60/250,077
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,327
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/250,176
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 10/023,617
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/49416
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/256,249
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/256,405
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 10/083,248
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: PCT/US01/46717
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/242,324
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,518
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/241,989
; PRIOR FILING DATE: 2000-10-20

```

```

; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-176-306-40

Query Match      88.2%; Score 30; DB 16; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 LKRLPK 7
Db      168 LKRLPK 173

RESULT 8
US-10-108-260A-3713
; Sequence 3713, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3713
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3713

Query Match      85.3%; Score 29; DB 12; Length 178;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLP 6
Db      139 SLKRLP 144

RESULT 9
US-09-815-242-5295
; Sequence 5295, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```


; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5295
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5295

Query Match 85.3%; Score 29; DB 9; Length 271;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 122 SLKELPK 128

RESULT 10
US-08-781-986A-5196
; Sequence 5196, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5196:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 273 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-781-986A-5196

Query Match 85.3%; Score 29; DB 8; Length 273;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 122 SLKELPK 128

RESULT 11
US-09-815-242-12548
; Sequence 12548, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12548
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12548

Query Match 85.3%; Score 29; DB 9; Length 273;
Best Local Similarity 85.7%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 122 SLKELPK 128

RESULT 12
US-10-100-121-20
; Sequence 20, Application US/10100121
; Publication No. US20030097686A1
; GENERAL INFORMATION:
; APPLICANT: Knauf, Vic C
; APPLICANT: Thompson, Gregory
; TITLE OF INVENTION: PLANT SEED OILS
; FILE REFERENCE: MTC 6823.1
; CURRENT APPLICATION NUMBER: US/10/100,121
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 08/926,522
; PRIOR FILING DATE: 1997-09-09
; PRIOR APPLICATION NUMBER: US 08/458,173
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: US 07/949,102
; PRIOR FILING DATE: 1992-09-21
; PRIOR APPLICATION NUMBER: US 07/762,762
; PRIOR FILING DATE: 1991-09-16
; PRIOR APPLICATION NUMBER: WO/US91/01746
; PRIOR FILING DATE: 1991-03-14
; PRIOR APPLICATION NUMBER: PCT/US91/05801
; PRIOR FILING DATE: 1991-08-15

; PRIOR APPLICATION NUMBER: US 07/615,784
; PRIOR FILING DATE: 1990-11-14
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Brassica campestris
US-10-100-121-20

Query Match 85.3%; Score 29; DB 15; Length 362;
Best Local Similarity 71.4%; Pred. No. 5.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 235 SLKRLPK 241

RESULT 13

US-09-853-625B-12
; Sequence 12, Application US/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION NUMBER: 09/053,587

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 374 AMINO ACIDS

; TYPE: AMINO ACID

; STRANDEDNESS: <Unknown>

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

; SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-853-625B-12

Query Match 85.3%; Score 29; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLP 6
|||:|
Db 207 SLKRLP 212

RESULT 14

US-09-901-910-7

; Sequence 7, Application US/09901910

; Publication No. US20030012768A1

; GENERAL INFORMATION:

; APPLICANT: Li, Haodong

; APPLICANT: Adams, Mark

; APPLICANT: Calenda Valerie

; TITLE OF INVENTION: Connective Tissue Growth Factor-2

; FILE REFERENCE: PFI26P2

; CURRENT APPLICATION NUMBER: US/09/901,910

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/348,815

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: 08/459,101

; PRIOR FILING DATE: 1995-06-02

; PRIOR APPLICATION NUMBER: PCT/US94/07736

; PRIOR FILING DATE: 1994-07-12

; PRIOR APPLICATION NUMBER: 60/217,402

; PRIOR FILING DATE: 2000-07-11

; PRIOR APPLICATION NUMBER: 60/291,642

; PRIOR FILING DATE: 2001-05-18

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 7

; LENGTH: 375

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-901-910-7

Query Match

85.3%; Score 29; DB 11; Length 375;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 SLKRLP 6

|||:|

Db 207 SLKRLP 212

RESULT 15

US-09-853-625B-11

; Sequence 11, Application US/09853625B

; Patent No. US20020049304A1

; GENERAL INFORMATION:

; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

; TITLE OF INVENTION: Human CCN-Like Growth Factor

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/853,625B

; FILING DATE: 14-May-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/053,587

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: MULLINS, J.G.

; REGISTRATION NUMBER: 33,073

; REFERENCE/DOCKET NUMBER: 325800-442

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 201-994-1700

```

; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 379 AMINO ACIDS
;   TYPE: AMINO ACID
;   STRANDEDNESS: <Unknown>
;   TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-853-625B-11

Query Match      85.3%; Score 29; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SLKRLP 6
      |||||
Db      202 SLKRLP 207
```

Search completed: February 11, 2004, 17:54:10
Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-7

Perfect score: 34

Sequence: 1 SLKRLPK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	100.0	3118	4	US-09-579-181-1
2	30	88.2	145	4	US-09-732-210-1428
3	29	85.3	139	4	US-09-252-991A-24681
4	29	85.3	319	1	US-07-745-206A-17
5	29	85.3	319	2	US-08-311-363-17
6	29	85.3	362	1	US-08-464-523B-32
7	29	85.3	362	2	US-08-948-176-25
8	29	85.3	362	2	US-08-440-845D-6
9	29	85.3	362	3	US-08-868-458-6
10	29	85.3	362	4	US-09-303-592-2
11	29	85.3	362	4	US-09-303-592-4
12	29	85.3	366	2	US-08-605-106-11
13	29	85.3	367	1	US-08-075-533-2
14	29	85.3	367	2	US-08-948-176-2
15	29	85.3	367	5	PCT-US91-09160-2
16	29	85.3	370	1	US-07-662-007B-37
17	29	85.3	370	1	US-07-824-247-37
18	29	85.3	370	3	US-08-470-204A-37
19	29	85.3	374	1	US-08-468-847B-12
20	29	85.3	375	2	US-08-459-101A-2
21	29	85.3	379	1	US-08-468-847B-11
22	29	85.3	379	4	US-09-142-569-2
23	29	85.3	381	4	US-09-142-569-4
24	29	85.3	381	4	US-09-348-815-2
25	29	85.3	382	1	US-07-824-247-42
26	29	85.3	382	1	US-08-142-473A-2
27	29	85.3	382	1	US-08-464-523B-26

28	29	85.3	382	1	US-08-469-203A-2	Sequence 2, Appli
29	29	85.3	382	1	US-08-469-203A-2	Sequence 2, Appli
30	29	85.3	382	2	US-08-605-106-12	Sequence 12, Appl
31	29	85.3	382	2	US-08-440-845D-1	Sequence 1, Appli
32	29	85.3	382	3	US-08-470-204A-42	Sequence 42, Appl
33	29	85.3	382	3	US-08-868-458-1	Sequence 1, Appli
34	29	85.3	591	4	US-09-252-991A-26716	Sequence 26716, A
35	29	85.3	1182	3	US-09-041-886-21	Sequence 21, Appl
36	29	85.3	1754	1	US-07-745-206A-13	Sequence 13, Appl
37	29	85.3	1754	2	US-08-311-363-13	Sequence 13, Appl
38	29	85.3	2237	1	US-08-455-543A-48	Sequence 48, Appl
39	29	85.3	2237	2	US-08-223-305C-48	Sequence 48, Appl
40	29	85.3	2237	4	US-09-268-163-8	Sequence 8, Appli
41	29	85.3	2265	2	US-08-149-097D-36	Sequence 36, Appl
42	29	85.3	2336	4	US-09-268-163-10	Sequence 10, Appl
43	29	85.3	2337	3	US-08-713-118-2	Sequence 2, Appli
44	29	85.3	2337	3	US-09-452-007-2	Sequence 2, Appli
45	29	85.3	2339	1	US-08-455-543A-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chrivia, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1

Query Match 100.0%; Score 34; DB 4; Length 3118;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||
Db 69 SLKRLPK 75

RESULT 2
US-09-732-210-1428
; Sequence 1428, Application US/09732210
; Patent No. 6573361
; GENERAL INFORMATION:
; APPLICANT: Bunkers, Greg J.
; APPLICANT: Liang, Jihong
; APPLICANT: Mittanck, Cindy A.
; APPLICANT: Seale, Jeffrey W.
; APPLICANT: Wu, Yonnie S.
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
; FILE REFERENCE: 38-21(15036)B
; CURRENT APPLICATION NUMBER: US/09/732,210
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,513
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: US 60/169,340
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 1753
; SEQ ID NO 1428

; LENGTH: 145
; TYPE: PRT
; ORGANISM: Xiphophorus maculatus
US-09-732-210-1428

Query Match 88.2%; Score 30; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7
Db 57 LKRLPK 62

RESULT 3
US-09-252-991A-24681
; Sequence 24681, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24681
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24681

Query Match 85.3%; Score 29; DB 4; Length 139;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 67 SLRLPK 73

RESULT 4
US-07-745-206A-17
; Sequence 17, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McCue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815

Query Match 85.3%; Score 29; DB 2; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-745-206A-17

Query Match 85.3%; Score 29; DB 1; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 23 TIKRLPK 29

RESULT 5
US-08-311-363-17
; Sequence 17, Application US/08311363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 319 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-311-363-17

Query Match 85.3%; Score 29; DB 2; Length 319;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 23 TIKRLPK 29

RESULT 6

US-08-464-523B-32
; Sequence 32, Application US/08464523B
; Patent No. 5723761
; GENERAL INFORMATION:
; APPLICANT: Toni A. Voelker
; APPLICANT: Ling Yuan
; APPLICANT: Jean Kridl
; APPLICANT: Deborah Hawkins
; APPLICANT: Aubrey Jones
; TITLE OF INVENTION: Plant Acyl ACP Thioesterase
; TITLE OF INVENTION: Sequences
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,523B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13131
; FILING DATE: 10-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,004
; FILING DATE: 10-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 100-1WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-523B-32

Query Match 85.3%; Score 29; DB 1; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
Db 235 SLKIPK 241

RESULT 7

US-08-948-176-25
; Sequence 25, Application US/08948176

; Patent No. 5945585
; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: YADAV, NARENDRA S.
; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES
; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT
; TITLE OF INVENTION: OIL COMPOSITION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,176
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/631,264
; FILING DATE: DECEMBER 20, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: CHRISTENEURY, LYNN M.
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: CR-8926-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-948-176-25

Query Match 85.3%; Score 29; DB 2; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
Db 235 SLKIPK 241

RESULT 8

US-08-440-845D-6
; Sequence 6, Application US/08440845D
; Patent No. 5955329
; GENERAL INFORMATION:
; APPLICANT: Yuan, L.
; APPLICANT: Kridl, J.
; APPLICANT: Dehesh, K.
; APPLICANT: Knauf, V.
; TITLE OF INVENTION: Engineering Plant Thioesterases For
; TITLE OF INVENTION: Altered Substrate Specificity.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
; COMPUTER: Apple Macintosh


```

; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,845D
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-845D-6

Query Match      85.3%; Score 29; DB 2; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 9
US-08-868-458-6
; Sequence 6, Application US/08868458
; Patent No. 6150512
; GENERAL INFORMATION:
; APPLICANT: Yuan, L.
; TITLE OF INVENTION: Engineering Plant Thioesterases And
; Disclosure of Plant Thioesterases
; Having No. 6150512e1 Substrate Specificity
;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: IBM PC
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: Microsoft Word For Window 95 7.0a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/868,458
; FILING DATE: 03-Jun-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07064
; FILING DATE: 15-MAY-96
; APPLICATION NUMBER: 08/537,083
; FILING DATE: 29-SEPT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 residues

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-868-458-6

Query Match      85.3%; Score 29; DB 3; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 10
US-09-303-592-2
; Sequence 2, Application US/09303592
; Patent No. 6426448
; GENERAL INFORMATION:
; APPLICANT: BOOTH, JOHN R.
; APPLICANT: BROGLIE, RICHARD M.
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: KINNEY, ANTHONY J.
; APPLICANT: KNOWLTON, SUSAN
; APPLICANT: SEBASTIAN, SCOTT A.
; APPLICANT: E.I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: NOVEL GENE COMBINATIONS THAT ALTER THE QUALITY AND
; TITLE OF INVENTION: FUNCTIONALITY OF SOYBEAN OIL
; FILE REFERENCE: BB-1156
; CURRENT APPLICATION NUMBER: US/09/303,592
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 60/085,030
; EARLIER FILING DATE: MAY 11, 1998
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 2
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Brassica napus
; US-09-303-592-2

Query Match      85.3%; Score 29; DB 4; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 11
US-09-303-592-4
; Sequence 4, Application US/09303592
; Patent No. 6426448
; GENERAL INFORMATION:
; APPLICANT: BOOTH, JOHN R.
; APPLICANT: BROGLIE, RICHARD M.
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: KINNEY, ANTHONY J.
; APPLICANT: KNOWLTON, SUSAN
; APPLICANT: SEBASTIAN, SCOTT A.
; APPLICANT: E.I. DU PONT DE NEMOURS AND COMPANY
; TITLE OF INVENTION: NOVEL GENE COMBINATIONS THAT ALTER THE QUALITY AND
; TITLE OF INVENTION: FUNCTIONALITY OF SOYBEAN OIL
; FILE REFERENCE: BB-1156
; CURRENT APPLICATION NUMBER: US/09/303,592
; CURRENT FILING DATE: 1999-05-03
; EARLIER APPLICATION NUMBER: 60/085,030
; EARLIER FILING DATE: MAY 11, 1998
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 4

```

```
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Brassica napus
US-09-303-592-4

Query Match      85.3%; Score 29; DB 4; Length 362;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      235 SLKKIPK 241

RESULT 12
US-08-605-106-11
; Sequence 11, Application US/08605106
; Patent No. 5910631
; GENERAL INFORMATION:
; APPLICANT: Topfer, R.
; APPLICANT: Martini, N.
; APPLICANT: Schell, J.
; TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,106
; FILING DATE: 23-SEPT-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02935
; FILING DATE: 01-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 235.001US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-106-11

Query Match      85.3%; Score 29; DB 2; Length 366;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      234 SLKKIPK 240

RESULT 13
US-08-075-533-2
; Sequence 2, Application US/08075533
; Patent No. 5530186
```

```
; GENERAL INFORMATION:
; APPLICANT: Hitz, William D.
; APPLICANT: Yadav, Narendra S.
; TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
; TITLE OF INVENTION: Thioesterase Genes
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. du Pont de Nemours and Company
; STREET: 1007 Market Street
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/075,533
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631,264
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Morrissey, Bruce W.
; REGISTRATION NUMBER: 30,663
; REFERENCE/DOCKET NUMBER: CR-8926-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (302) 992-4927
; TELEFAX: (302) 892-7949
; TELEX: 835420
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 367 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-075-533-2

Query Match      85.3%; Score 29; DB 1; Length 367;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 SLKRLPK 7
Db      232 SLKKIPK 238

RESULT 14
US-08-948-176-2
; Sequence 2, Application US/08948176
; Patent No. 5945585
; GENERAL INFORMATION:
; APPLICANT: HITZ, WILLIAM D.
; APPLICANT: YADAV, NARENDRA S.
; TITLE OF INVENTION: ACYL-ACP THIOESTERASES GENES
; TITLE OF INVENTION: AND THEIR USE IN ALTERING PLANT
; TITLE OF INVENTION: OIL COMPOSITION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0
```

PCT-US91-09160-2

Query Match	85.3%	Score 29;	DB 5;	Length 367;
Best Local Similarity	71.4%	Pred. No.	1.5e+02;	
Matches	5;	Conservative	2;	Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 232 SLKKIPK 238

Search completed: February 11, 2004, 17:13:37
Job time : 11.4167 secs

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,176
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA: 07/631,264
APPLICATION NUMBER: 07/631,264
FILING DATE: DECEMBER 20, 1990
ATTORNEY/AGENT INFORMATION:
NAME: CHRISTENBURY, LYNNE M.
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: CR-8926-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 367 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-948-176-2

Query Match	85.3%	Score 29;	DB 2;	Length 367;
Best Local Similarity	71.4%	Pred. No.	1.5e+02;	
Matches	5;	Conservative	2;	Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
|||:|
Db 232 SLKKIPK 238

RESULT 15
PCT-US91-09160-2
Sequence 2, Application PC/TUS9109160
GENERAL INFORMATION:
APPLICANT: Hitz, William D.
APPLICANT: Yadav, Narendra S.
TITLE OF INVENTION: Nucleotide Sequences of Soybean Acyl-ACP
TITLE OF INVENTION: Thioesterase Genes
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. du Pont de Nemours and Company
STREET: 1007 Market Street
CITY: Wilmington
STATE: Delaware
COUNTRY: U.S.A
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/09160
FILING DATE: 19911216
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/631,264
FILING DATE: 20-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Morrissey, Bruce W.
REGISTRATION NUMBER: 30,663
REFERENCE/DOCKET NUMBER: CR-8926-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (302) 992-4927
TELEFAX: (302) 892-7949
TELEX: 835420
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 367 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-7

Perfect score: 34

Sequence: 1 SLKRLPX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	7	AAE14552	Human alpha-synuc
2	34	100.0	3118	AAB50362	Human SRCAP, Homo
3	32	94.1	478	ABB53479	Lactococcus lactis
4	31	91.2	45	ABB95861	Human testicular a
5	31	91.2	45	AAM95160	Human reproductive
6	30	88.2	62	ABB03831	Human musculoskele
7	30	88.2	62	ABU13125	Novel human muscul
8	30	88.2	219	AAU99919	Human 49875 DEAD t
9	30	88.2	254	ABP78272	N. gonorrhoeae ami

10	30	88.2	396	23	ABB49928	Listeria monocytog
11	30	88.2	448	22	ABG24660	Novel human diagno
12	30	88.2	966	22	ABB68113	Drosophila melanog
13	29	85.3	102	23	ABP09012	Human ORFX protein
14	29	85.3	165	21	AAG06433	Arabidopsis thalia
15	29	85.3	165	23	ABG60070	Human DITHP polype
16	29	85.3	171	21	AAG54995	Arabidopsis thalia
17	29	85.3	174	21	AAG14734	Arabidopsis thalia
18	29	85.3	217	21	AAG06432	Arabidopsis thalia
19	29	85.3	229	21	AAG06431	Arabidopsis thalia
20	29	85.3	229	21	AAG39849	Arabidopsis thalia
21	29	85.3	230	21	AAG14733	Arabidopsis thalia
22	29	85.3	235	21	AAG39986	Arabidopsis thalia
23	29	85.3	271	22	AAU33799	Staphylococcus aur
24	29	85.3	273	18	AAW89748	Staphylococcus aur
25	29	85.3	273	22	AAU36955	Staphylococcus thalia
26	29	85.3	291	21	AAG39985	Arabidopsis thalia
27	29	85.3	306	21	AAG14732	Arabidopsis thalia
28	29	85.3	362	13	AAR29172	Brassica thioester
29	29	85.3	362	15	AAR54948	Brassica campestri
30	29	85.3	362	16	AAR74150	Brassica acyl thio
31	29	85.3	362	19	AAW44333	Brassica campestri
32	29	85.3	362	20	AAU28638	Brassica napus acy
33	29	85.3	362	20	AAW87467	Brassica rapa Cl8:
34	29	85.3	362	21	AAG39848	Arabidopsis thalia
35	29	85.3	362	21	AAU53760	An acyl-ACP thioes
36	29	85.3	362	21	AAU53761	An acyl-ACP thioes
37	29	85.3	362	23	ABG97569	Acyl acyl carrier
38	29	85.3	362	23	ABB92431	Herbicidally activ
39	29	85.3	366	14	AAR41673	Rape acyl-ACP thio
40	29	85.3	366	14	AAR41674	Rape acyl-ACP thio
41	29	85.3	367	13	AAR28339	Acyl-ACP thioester
42	29	85.3	367	20	AAU28632	Soybean seed acyl-
43	29	85.3	367	21	AAG39984	Arabidopsis thalia
44	29	85.3	367	23	ABB92915	Herbicidally activ
45	29	85.3	370	12	AAR14802	Bay thioesterase.

ALIGNMENTS

RESULT 1

AAE14552

ID AAE14552 standard; peptide; 7 AA.

XX

AC AAE14552;

XX

DT 17-MAY-2002 (first entry)

XX

DE Human alpha-synuclein aggregation inhibitor #7.

XX

KW Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;

KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;

KW multiple system atrophy; Hallervorden-Spatz disease; human.

XX

OS Homo sapiens.

XX

PN WO200204482-A1

XX

PD 17-JAN-2002.

XX

PF 06-JUL-2001; 2001WO-US21379.

XX

PR 07-JUL-2000; 2000US-217319P.

PR 28-MAR-2001; 2001US-279199P.

XX

PA (PANA-) PANACEA PHARM INC.

XX

PI Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX

DR WPI; 2002-179695/23.

XX

PT Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
PS Claim 40; Page 37; 52pp; English.
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 34; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
DB 1 SLKRLPK 7
RESULT 2
AAB50362
ID AAB50362 standard; protein; 3118 AA.
AC AAB50362;
XX 12-MAR-2001 (first entry)
DT Human SRCAP.
DE
XX Human; SRCAP; Snf2 related CBP activator protein; antiviral; CREB;
KW cAMP regulatory element; CREB binding protein; CBP; ATPase;
KW transcription activation; DEAD box RNA dependent helicase;
KW adenoviral DBP protein; beta-actin; nuclear receptor; viral infection.
XX Homo sapiens.
OS
XX WO200073467-A1.
PN
XX 07-DEC-2000.
PD
XX 25-MAY-2000; 2000WO-US14719.
PF
XX 27-MAY-1999; 99US-0136620.
PR 25-MAY-2000; 2000US-0579181.
XX
XX (UYSL-) UNIV SAINT LOUIS.
PA
XX Chrivia J, Yaciuk P;
PI
XX WPI; 2001-061545/07.
DR
DR N-PSDB; AAC89859.
XX
XX Snf2 related cAMP regulatory element (CREB) binding protein (CBP)
PT activator protein, capable of co-activating CREB binding protein,
PT useful for modulating transcription and for affecting viral infection -
XX
PS Claim 6; Page 77-86; 103pp; English.
XX
XX The present sequence is an Snf2 related CREB (cAMP regulatory element)
CC binding protein (CBP) activator protein (SRCAP) polypeptide. It has
CC ATPase activity and is capable of activating transcription. SRCAP
CC polypeptides are useful for activating transcription in a cell, for
CC enhancing CREB (cAMP regulatory element) binding protein (CBP)-mediated
CC activation of transcription in a cell, for treating a patient having a
CC disease involving a function such as insufficient transcription of a

CC gene, e.g. a gene mediated by CBP co-activation, DEAD box RNA dependent
CC helicase, adenoviral DBP protein, beta-actin or a nuclear receptor
CC affected by SRCAP protein. Compounds that modulate SRCAP function, such
CC as antibodies, antisense molecules, polynucleotides or ribozymes, are
CC useful for treating diseases mediated by SRCAP-activated transcription,
CC for example, infection by adenovirus, hepatitis C virus, human
CC immunodeficiency virus type-1, Epstein Barr virus, cytomegalovirus or
CC hepatitis B virus.
XX Sequence 3118 AA;
SQ
Query Match 100.0%; Score 34; DB 22; Length 3118;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
DB 69 SLKRLPK 75
RESULT 3
ABB53479
ID ABB53479 standard; Protein; 478 AA.
XX ABB53479;
AC
XX 16-MAY-2002 (first entry)
DT
XX Lactococcus lactis protein bgls.
DE
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
KW Lactococcus lactis IL1403.
XX
OS FR2807446-A1.
PN
XX 12-OCT-2001.
PD
XX 11-APR-2000; 2000FR-0004630.
PF
XX 11-APR-2000; 2000FR-0004630.
PR
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX WPI; 2002-043418/06.
DR
XX New nucleotide sequence useful in the identification of Lactococcus
PT lactis and related species -
XX
PS Claim 6; SEQ ID No 181; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 478 AA;
Query Match 94.1%; Score 32; DB 23; Length 478;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
DB 1 SLKRLPK 75

Db 455 SLKRIPK 461

RESULT 4

ABB95861

ID ABB95861 standard; Protein; 45 AA.

XX AC ABB95861;

XX DT 21-JUN-2002 (first entry)

XX DE Human testicular antigen SEQ ID NO: 1245.

XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;

XX KW reproductive system disorder; urinary system disorder; gene therapy;

XX KW cardiovascular disorder; respiratory disorder; neurological disorder;

XX KW gastrointestinal disease; infection; cytostatic.

XX OS Homo sapiens.

XX PN WO200155317-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01329.

XX PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 08-NOV-2000; 2000US-0246613.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.

PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.

PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.

PR 17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.

PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.


```
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483232/52.
XX
PT Nucleic acids encoding 973 human testicular antigen polypeptides,
PT useful for preventing, diagnosing and/or treating testicular cancer
XX
PS Claim 11; SEQ ID NO 1245; 766pp; English.
XX
CC The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a protein of the
CC invention.
XX
SQ Sequence 45 AA;
Query Match 91.2%; Score 31; DB 22; Length 45;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 SLKRLPK 7
Db 29 ALKRLPK 35
RESULT 5
AAM95160
ID AAM95160 standard; Protein; 45 AA.
XX
AC AAM95160;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen SEQ ID NO: 3818.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
```


PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-451937/48.

N-PSDB; AAL35413.

Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -

Claim 11; SEQ ID NO 1778; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 62 AA;

Query Match 88.2%; Score 30; DB 22; Length 62;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7

|||||

Db 20 LKRLPK 25

Db 20 LKRLPK 25

RESULT 8
AAU99919
ID AAU99919 standard; Protein; 219 AA.
XX
AC AAU99919;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human 49875 DEAD type helicase domain consensus sequence (SMART).
XX
KW 49875; DEAD type helicase domain; infection;
KW haematopoietic disorder; blood clotting disorder; cancer;
KW autoimmune disorder; leukaemia; immunological disorder;
KW cardiovascular disorder; neurological disorder; cellular proliferation;
KW red blood cell disorder; viral disease; neurological disorder.
XX
OS Synthetic.
XX
PN WO200240656-A2.
XX
PD 23-MAY-2002.
XX
PF 14-NOV-2001; 2001WO-US45291.
XX
PR 14-NOV-2000; 2000US-248331P.
PR 14-NOV-2000; 2000US-248362P.
PR 14-NOV-2000; 2000US-248365P.
PR 30-NOV-2000; 2000US-250077P.
PR 30-NOV-2000; 2000US-250176P.
PR 30-NOV-2000; 2000US-250327P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers RE, Curtis RAJ, Glucksmann MA;
XX
XX WPI; 2002-508325/54.
DR
XX Isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041
PT or 84234 polypeptides, useful as reagents or targets for treating or
PT diagnosing pain or metabolic, liver, kidney, or cardiovascular
PT disorders -
XX
PS Disclosure; Fig 8; 298pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel
CC isolated 47476, 67210, 49875, 46842, 33201, 83378, 84233, 64708, 85041
CC or 84234 proteins. The method of the invention is useful for treating a
CC disorder characterised by aberrant activity of 47476, 67210, 49875,
CC 46842, 33201, 83378, 84233, 64708, 85041 or 84234-expressing cell in a
CC subject. The protein molecules can act as novel diagnostic targets and
CC therapeutic agents for controlling aberrant or deficient signal
CC transduction resulting, in e.g., haematopoietic disorders, including
CC blood clotting disorders, autoimmune disorders, or disorders related to
CC an inability to clear infections (e.g., viral or bacterial infections),
CC as well as disorders related to abnormal cellular proliferation or
CC differentiation, e.g., leukaemia. They may also be used to control
CC disorders of metabolic imbalance (e.g., disorders of lipopolysaccharide
CC biosynthesis or glycogen synthesis) immunological disorders,
CC cardiovascular disorders, neurological disorders, or cellular
CC proliferation and/or differentiation disorders, e.g., cancer, cell
CC motility and adhesion disorders differentiative disorders, red blood
CC cell disorders, viral diseases, neurological disorders (e.g., brain
CC disorders), pain or metabolic disorders, liver disorders, kidney
CC disorders, disorders of the small intestine, disorders of metal ion
CC imbalance, protein trafficking disorders and disorders associated with
CC bone metabolism. The sequences of the invention are also useful for
CC screening assays, predictive medicine (e.g., diagnostic assays,
CC prognostic assays, monitoring clinical trials, and pharmacogenetics);
CC and methods of treatment (e.g., therapeutic and prophylactic). The
CC present sequence represents a predicted consensus sequence motif

CC found in the protein of the invention.
XX
SQ Sequence 219 AA;
Query Match 88.2%; Score 30; DB 23; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LKRLPK 7
Db 168 LKRLPK 173
RESULT 9
ABP78272
ID ABP78272 standard; Protein; 254 AA.
XX
AC ABP78272;
XX
DT 07-MAR-2003 (first entry)
XX
DE N. gonorrhoeae amino acid sequence SEQ ID 3074.
XX
KW Antibacterial; infection; vaccine; gene therapy.
XX
OS Neisseria gonorrhoeae.
XX
PN WO200279243-A2.
XX
PD 10-OCT-2002.
XX
PF 12-FEB-2002; 2002WO-IB02069.
XX
PR 12-FEB-2001; 2001GB-0003424.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Fontana MR, Pizza M, Massignani V, Monaci E;
XX
XX WPI; 2003-058415/05.
DR N-PSDB; ABZ39242.
DR
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection -
PT
XX Disclosure; Page 412; 815pp; English.
PS
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention.
XX
SQ Sequence 254 AA;
Query Match 88.2%; Score 30; DB 24; Length 254;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
Db 45 SLKRLPE 51
RESULT 10
ABB49928
ID ABB49928 standard; Protein; 396 AA.
XX
AC ABB49928;
XX

DT 05-FEB-2002 (first entry)
XX Listeria monocytogenes protein #2632.
DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX vitamin B12; bacterial infection; disease.
KW Listeria monocytogenes.
XX WO200177335-A2.
XX 18-OCT-2001.
XX 11-APR-2001; 2001WO-FR01118.
XX 11-APR-2000; 2000FR-0004629.
XX (INSP) INST PASTEUR.
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Krefit J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baguero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX WPI; 2002-010914/01.
DR Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and
PT related polypeptides -
PT Claim 6; SEQ ID No 2633; 192pp; French.
XX The present invention relates to the genome sequence of Listeria
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
CC it are useful for selecting probes and primers for detecting genes in L.
CC monocytogenes and related organisms, and for studying genetic
CC polymorphisms and other genomes. The present invention is a protein
CC encoded by the genome sequence of the present invention. Proteins
CC expressed from the genome sequence are useful for raising specific
CC antibodies, identification of L. monocytogenes and related organisms, and
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
CC B12. The genome sequence and proteins encoded by it are also useful for
CC selecting compounds that regulate gene expression and cell replication
CC and modulate L. monocytogenes-related diseases. In addition, the genome
CC sequence and proteins encoded by it are useful in pharmaceutical and
CC vaccines compositions for the treatment or prevention of infections by L.
CC monocytogenes and related organisms.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 396 AA;
Query Match 88.2%; Score 30; DB 23; Length 396;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SLKRLPK 7
| | | | |
Db 181 SLKRLPE 187

RESULT 11
ABG24660
ID ABG24660 standard; Protein; 448 AA.
XX
AC ABG24660;
XX 18-FEB-2002 (first entry)
DT

XX DE Novel human diagnostic protein #24651.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
OS WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS88847.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX Claim 20; SEQ ID No 55019; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 448 AA;

Query Match 88.2%; Score 30; DB 22; Length 448;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LKRLPK 7
| | | | |
Db 134 LKRLPK 139

RESULT 12
ABB68113
ID ABB68113 standard; Protein; 966 AA.
XX
AC ABB68113;
XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 31131.
DE

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX 23-MAR-2000; 2000US-191637P.
PR
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEXE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR
DR N-PSDB; ABL12216.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 31131; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 966 AA;
Query Match 88.2%; Score 30; DB 22; Length 966;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
Db 228 SLKQLPK 234
RESULT 13
ABP09012
ID ABP09012 standard; Protein; 102 AA.
XX
AC ABP09012;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX protein sequence SEQ ID NO:18006.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.

XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US10836.
PF
XX 30-MAY-2000; 2000US-206132P.
PR
PR 29-AUG-2000; 2000US-228716P.
XX
XX (CURA-) CURAGEN CORP.
PA
XX Shinkets RA, Leach MD;
PI
XX WPI; 2002-106308/14.
DR
DR N-PSDB; ABL24764.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -
XX
PS Disclosure; SEQ ID 18006; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABL15762 to ABL27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 102 AA;
Query Match 85.3%; Score 29; DB 23; Length 102;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SLKRLPK 7
Db 65 SIKKLPK 71
RESULT 14
AAG06433
ID AAG06433 standard; Protein; 165 AA.
XX
XX AAG06433;
AC
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 3204.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.

us-09-901-187c-7.rag

Thu Feb 12 09:07:48 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- SPTREMBL 23:**
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	100.0	483	5 Q95SP1	Q95sp1 drosophila
2	36	100.0	533	5 Q9VHV3	Q9vhv3 drosophila
3	33	91.7	812	11 Q35483	O35483 mus musculu
4	32	88.9	438	2 Q8KZP3	Q8kzp3 thermus the
5	31	86.1	75	16 Q8Y8F8	Q8y8f8 listeria mo
6	31	86.1	145	16 Q8YP84	Q8yp84 anabaena sp
7	31	86.1	153	16 Q8DKN7	Q8dkn7 synechococc
8	31	86.1	195	10 Q9C9A5	Q9c9a5 arabidopsis
9	31	86.1	213	10 Q39156	Q39156 arabidopsis
10	31	86.1	280	16 Q8Y4N0	Q8y4n0 listeria mo
11	31	86.1	281	16 Q928M7	Q928m7 listeria in
12	31	86.1	358	2 Q8GP85	Q8gp85 streptococc
13	31	86.1	384	5 Q9V629	Q9v629 drosophila
14	31	86.1	386	16 Q8YWS7	Q8yws7 anabaena sp
15	31	86.1	391	2 Q9RMT7	Q9rmt7 streptomyce
16	31	86.1	415	16 Q9FC33	Q9fc33 streptomyce

17	31	86.1	491	5 Q9U622	Q9u622 drosophila
18	31	86.1	493	2 Q8GMH4	Q8gmh4 streptomyce
19	31	86.1	1747	10 Q8LHA4	Q8lha4 oryza sativ
20	30	83.3	84	2 Q8KN68	Q8kn68 pseudomonas
21	30	83.3	273	16 Q8PDE6	Q8pde6 xanthomonas
22	30	83.3	329	16 Q8P4X4	Q8p4x4 xanthomonas
23	30	83.3	404	16 Q9Z3E5	Q9z3e5 xanthomonas
24	30	83.3	404	16 Q87641	Q87641 xanthomonas
25	30	83.3	404	16 Q8PL87	Q8pl87 xanthomonas
26	30	83.3	404	16 Q8PL76	Q8pl76 xanthomonas
27	30	83.3	407	16 Q8PRM5	Q8prm5 xanthomonas
28	30	83.3	412	5 Q962C3	Q962c3 caenorhabdi
29	30	83.3	440	5 Q19730	Q19730 caenorhabdi
30	30	83.3	449	8 Q9TM09	Q9tm09 cyanidium c
31	30	83.3	607	10 Q9FL24	Q9fl24 arabidopsis
32	30	83.3	1254	5 Q9U308	Q9u308 caenorhabdi
33	29	80.6	40	16 Q989U1	Q989u1 rhizobium l
34	29	80.6	165	10 Q9XHP7	Q9xhp7 ceratopteri
35	29	80.6	170	16 Q8YN03	Q8yn03 anabaena sp
36	29	80.6	184	2 Q45156	Q45156 bacillus ps
37	29	80.6	203	16 Q9X109	Q9x109 thermotoga
38	29	80.6	210	10 Q9SE48	Q9se48 oryza sativ
39	29	80.6	210	10 Q9SE47	Q9se47 oryza sativ
40	29	80.6	216	11 Q8CB99	Q8cb99 mus musculu
41	29	80.6	231	10 Q8LHJ7	Q8lhj7 oryza sativ
42	29	80.6	245	16 Q8CKZ5	Q8ckz5 yersinia pe
43	29	80.6	251	16 Q8ZFG2	Q8zfg2 yersinia pe
44	29	80.6	255	11 Q9D2K3	Q9d2k3 mus musculu
45	29	80.6	274	16 Q8Y1F3	Q8ylf3 ralstonia s

ALIGNMENTS

RESULT 1
Q95SP1 ID Q95SP1 PRELIMINARY; PRT; 483 AA.
AC Q95SP1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE GH11496p.
GN CG7800.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060678; AAL28226.1; --
DR FlyBase; FBgn0037552; CG7800.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 6.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
DR PROSITE; PS05056; LRR_TYPICAL; 1.
SQ SEQUENCE 483 AA; 54650 MW; 459439E8EFE17F44 CRC64;

Query Match 100.0%; Score 36; DB 5; Length 483;
Best Local Similarity 100.0%; Pred.No.13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||||

Db 453 RLRGRNQ 459

RESULT 2

Q9VHV3 PRELIMINARY; PRT; 533 AA.
AC Q9VHV3; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE CG7800 protein.
GN CG7800.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M.D., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003678; AAF54196.2; -.
DR FlyBase; FBgn0037552; CG7800.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR; 5.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 1.
DR PROSITE; PS0506; LRR_TYPICAL; 1.
SQ SEQUENCE 533 AA; 60296 MW; 3532E77037F67A6E CRC64;

Query Match 100.0%; Score 36; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 503 RLRGRNQ 509

RESULT 3

O35483 PRELIMINARY; PRT; 812 AA.
AC O35483;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Kruppel-related zinc finger protein.
GN ZFP316 OR EMZF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Yolk sac;
RA Rothstein J.L., Brafford P.;
RT "Murine Emz1 gene: Expression During Development.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031955; AAB87452.1; -.
DR HSSP; P08045; 1ZNF.
DR MGD; MGI:1860402; Zfp316.
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 14.
DR ProDom; PD000003; Znf_C2H2; 6.
DR SMART; SM00355; Znf_C2H2; 14.
DR PROSITE; PS00402; BPD_TRANS INN MEMBER; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 12.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 14.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 812 AA; 88817 MW; 8F89DE55FD063DCB CRC64;

Query Match 91.7%; Score 33; DB 11; Length 812;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 722 RVRGRNQ 728

RESULT 4

Q8KZP3 PRELIMINARY; PRT; 438 AA.
AC Q8KZP3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Preprotein translocase Secy subunit.
GN SECY.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8;
RA Masui R., Inoue Y., Shibata T., Miki K., Yokoyama S., Kuramitsu S.;
RT "preprotein translocase SecE subunit.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086887; BAC01134.1; -.
DR InterPro; IPR002208; Secy.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00344; secy; 1.
DR PRINTS; PR00303; SECYTRNLCASE.
DR TIGRFAMs; TIGR00967; 3a0501s007; 1.
SQ SEQUENCE 438 AA; 50075 MW; 3532E77037F67A6E CRC64;

Query Match 100.0%; Score 36; DB 5; Length 533;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
SQ SEQUENCE 438 AA; 48204 MW; 71BBEF926123A69E CRC64;

Query Match 88.9%; Score 32; DB 2; Length 438;
Best Local Similarity 85.7%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|||||
Db 432 RLRGRNR 438

RESULT 5
Q8Y8F8 PRELIMINARY; PRT; 75 AA.
ID Q8Y8F8
AC Q8Y8F8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein lmo0946.
GN LMO0946.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-Lopez N., Hain T., Hauf J., Jackson D.,
RA Gautier L., Goebel W., Gomez-Lopez N., Kneft J., Kuhn M., Kunst F., Kurapkat G.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL; AL591977; CAC99024.1; -.
DR Listlist; LMO00946; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 75 AA; 8810 MW; 8242CB314143B336 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
|||||
Db 70 RLRGRN 75

RESULT 6
Q8YP84 PRELIMINARY; PRT; 145 AA.
ID Q8YP84
AC Q8YP84;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein Ali4315.
GN Ali4315.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,

RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003596; BAB76014.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 145 AA; 15732 MW; D16371D9C4EBA237 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 145;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
|||||
Db 99 RLRGRN 104

RESULT 7
Q8DKN7 PRELIMINARY; PRT; 153 AA.
ID Q8DKN7
AC Q8DKN7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tlr0822 protein.
GN TLR0822.
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1."
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005371; BAC08373.1; -.
KW Complete proteome.
SQ SEQUENCE 153 AA; 16164 MW; 01B4E4268EDBFC90 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 153;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
|||||
Db 44 RLRGRN 49

RESULT 8
Q9C9A5 PRELIMINARY; PRT; 195 AA.
ID Q9C9A5
AC Q9C9A5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative transcription factor (Atlg1030/F23N20_2).
GN F23N20.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,

RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Dunn M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana";
RL Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,
RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,
RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cdna clones";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
RA Chung M.K., Kim C., Lin J., Liu S.X., Pham P.K., Sakano H., Shinn P.,
RA Yamada K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis ORF clones";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
annotation";
RL Genome Biol. 0:0-0(2002).
RN [6]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
DR EMBL; AC016972; AAG51688.1; -
DR EMBL; AY048231; AAK82494.1; -
DR EMBL; AF370609; AAK43928.1; -
DR EMBL; AY091697; AAM10296.1; -
DR EMBL; AY087093; AAM64653.1; -
DR HSSP; P06876; IMEK.
DR InterPro; IPR001005; Myb DNA binding.
DR Pfam; PF00249; myb DNA-binding; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 1.
SQ SEQUENCE 195 AA; 22098 MW; B270E3414A015967 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLRGRN 6
DB 22 RLRGRN 27
RESULT 9
Q39156 PRELIMINARY; PRT; 213 AA.
ID Q39156
AC Q39156;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Myb-related transcription factor.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Siliques;
RX MEDLINE=97149286; PubMed=8996094;
RA Kirik V., Baumlein H.;
RT "A novel leaf-specific myb-related protein with a single binding
repeat";
RL Gene 183:109-113(1996).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL; Z68157; CAA92280.1; -
DR HSSP; P06876; IMBK.
DR InterPro; IPR001005; Myb DNA binding.
DR Pfam; PF00249; myb DNA-binding; 1.
DR SMART; SM00717; SANT; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 213 AA; 24693 MW; 9DDB050F5FE40ADC CRC64;
Query Match 86.1%; Score 31; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLRGRN 6
DB 22 RLRGRN 27
RESULT 10
Q8Y4N0 PRELIMINARY; PRT; 280 AA.
ID Q8Y4N0
AC Q8Y4N0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein lmo2406.
GN LMO2406.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,

```
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitounnam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591983; CAD00484.1; -.
DR ListiList; LMO02406; -.
DR InterPro; IPR002763; DUF72.
DR Pfam; PF01904; DUF72; 1.
DR ProDom; PD117970; DUF72; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 280 AA; 32186 MW; 7C154678230D125C CRC64;

Query Match      86.1%; Score 31; DB 16; Length 280;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLGRNQ 7
Db      192 RLGRNQ 198

RESULT 11
Q928M7
ID Q928M7 PRELIMINARY; PRT; 281 AA.
AC Q928M7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein lin2505.
GN LIN2505.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitounnam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596172; CAC97732.1; -.
DR ListiList; LIN02505; -.
DR InterPro; IPR002763; DUF72.
DR Pfam; PF01904; DUF72; 1.
DR ProDom; PD117970; DUF72; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 32372 MW; 5FCBACFF0ECC6D87 CRC64;

Query Match      86.1%; Score 31; DB 16; Length 281;
Best Local Similarity 85.7%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 RLGRNQ 7
Db      193 RLGRNQ 199

RESULT 12
Q8GP85
```

```
ID Q8GP85 PRELIMINARY; PRT; 358 AA.
AC Q8GP85;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Eps9L.
GN Eps9L.
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE FROM N.A.
RA Rallu F., Ehrlich D.S., Renault P.,
RT "Diversity of eps operons in Streptococcus thermophilus.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF454499; AAN63755.1; -.
SQ SEQUENCE 358 AA; 40644 MW; 964CA552A8B16D99 CRC64;

Query Match      86.1%; Score 31; DB 2; Length 358;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLGRNQ 7
Db      60 RLGRNE 66

RESULT 13
Q9V629
ID Q9V629 PRELIMINARY; PRT; 384 AA.
AC Q9V629;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG13190 protein.
GN CG13190.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```

```
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AF003824; AAF58610.1; -.
DR FlyBase; FBgn0033662; CG13190.
SQ SEQUENCE 384 AA; 44833 MW; 0FF9AC48BCCB1AEC CRC64;

Query Match 86.1%; Score 31; DB 5; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRGRNQ 7
Db 347 LRGRNQ 352
|||||

RESULT 14
Q8YWS7 PRELIMINARY; PRT; 386 AA.
AC Q8YWS7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein All1523.
GN All1523.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003586; BAB7889.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 386 AA; 41303 MW; 23B333C3B98B5AF7 CRC64;

Query Match 86.1%; Score 31; DB 16; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
Db 299 RLGRN 304
|||||

RESULT 15
Q9RMT7 PRELIMINARY; PRT; 391 AA.
AC Q9RMT7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aminase.
GN PABB.
OS Streptomyces venezuelae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=54571;
```

```
RN SEQUENCE FROM N.A.
RP STRAIN=ISP5230;
RA Chang Z., Sun Y., Vining L.C.;
RT "pabA and pabB, the second set of p-aminobenzoic acid synthase genes
in Streptomyces venezuelae ISP5230.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF189258; AAF01062.1; -.
DR HSSP; Q06128; 1QDL.
DR InterPro; IPR005801; Anth_synth_chor.
DR InterPro; IPR006056; YjgF-like.
DR Pfam; PF00425; chorismate_bind; 1.
DR PRINTS; PR00095; ANTSNTHASE1.
DR PRODOM; PD000779; Anth_synth_chor; 1.
DR PROSITE; PS01094; UPE0076; 1.
SQ SEQUENCE 391 AA; 43299 MW; 64D3207A87B41065 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
Db 127 RLGRN 132
|||||

Search completed: February 11, 2004, 17:09:32
Job time : 26.5833 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	86.1	185	1 KGUA_FUSNN	Q8rh19 fusobacteri
2	31	86.1	207	1 KGUA_ECO57	Q8xd88 escherichia
3	30	83.3	157	1 YH13_STRPN	P58260 streptococc
4	30	83.3	207	1 KGUA_ECOLI	P24234 escherichia
5	30	83.3	207	1 KGUA_SALTI	Q8z2h9 salmonella
6	30	83.3	207	1 KGUA_SALTY	Q9x6m5 salmonella
7	30	83.3	207	1 KGUA_YERPE	Q8zjq2 yersinia pe
8	30	83.3	576	1 DNAB_MYCIT	Q9f5p4 mycobacteri
9	29	80.6	202	1 UMPK_ARATH	O04905 arabidopsis
10	29	80.6	312	1 TF2B_ORISA	Q8w0w3 oryza sativ
11	29	80.6	517	1 LEGA_PEA	P02857 pisum sativ
12	29	80.6	520	1 LEG2_PEA	P15838 pisum sativ
13	29	80.6	674	1 COAA_BACTJ	O87905 bacillus th
14	28	77.8	126	1 CV02_HUMAN	Q9y3m2 homo sapien
15	28	77.8	127	1 CV02_MOUSE	Q9dlc2 mus musculu
16	28	77.8	178	1 YK01_ARCFU	O28278 archaeoglob
17	28	77.8	199	1 KGUA_ANASP	Q8z0i7 anabaena sp
18	28	77.8	242	1 RNPH_NEIMA	Q9jtn2 neisseria m
19	28	77.8	242	1 RNPH_NEIMB	Q9jyn5 neisseria m
20	28	77.8	260	1 DCK_HUMAN	P27707 homo sapien
21	28	77.8	260	1 DCK_MOUSE	P43346 mus musculu
22	28	77.8	386	1 NESG_HUMAN	Q9ul16 homo sapien
23	28	77.8	395	1 PELO_DROME	P48612 drosophila
24	28	77.8	445	1 MEK1_SCHPO	Q10292 schizosacch
25	28	77.8	476	1 YHGF_NEIGO	Q51062 neisseria g
26	28	77.8	509	1 AGP4_MOUSE	Q9wvh6 mus musculu
27	28	77.8	517	1 Y967_TREPA	O83933 treponema p
28	28	77.8	676	1 HMUR_YERPE	Q56989 yersinia pe
29	28	77.8	687	1 HEMR_YEREN	P31499 yersinia en
30	28	77.8	757	1 YHGF_NEIMA	P57072 neisseria m
31	28	77.8	757	1 YHGF_NEIMB	Q51152 neisseria m
32	28	77.8	917	1 SMOO_HUMAN	P53814 homo sapien
33	28	77.8	1038	1 CIN8_YEAST	P27895 saccharomyc

34	28	77.8	1184	1 BIMC_EMENI	P17120 emericella
35	28	77.8	1714	1 YP93_CAEEL	Q09475 caenorhabdi
36	28	77.8	1816	1 LMA4_MOUSE	P97927 mus musculu
37	27	75.0	133	1 TOX5_BORPE	P04981 bordetella
38	27	75.0	204	1 VN5C_SENDF	P14253 sendai viru
39	27	75.0	204	1 VN5C_SENDF	P14254 sendai viru
40	27	75.0	204	1 VN5C_SENDF	P04861 sendai viru
41	27	75.0	204	1 VN5C_SENDF	P04862 sendai viru
42	27	75.0	209	1 SEGD_BPT4	P19895 bacterioph
43	27	75.0	235	1 GAG_MLVAB	P03333 abelson mur
44	27	75.0	239	1 RNPH_RICCN	Q92gz7 rickettsia
45	27	75.0	240	1 RNPH_RICPR	Q9zct5 rickettsia

ALIGNMENTS

RESULT 1					
KGUA_FUSNN					
ID	KGUA_FUSNN	STANDARD;	PRT;	185 AA.	
AC	Q8RH19;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Guanylate kinase (EC 2.7.4.8) (GMP kinase).				
GN	GMK OR FN2033.				
OS	Fusobacterium nucleatum (subsp. nucleatum).				
OC	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;				
OC	Fusobacterium.				
OX	NCBI_TaxID=76856;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 25586;				
RX	MEDLINE=21886394; PubMed=11889109;				
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,				
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,				
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,				
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,				
RA	Fonstein M., Kyrpides N., Overbeek R.;				
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium				
RT	nucleatum strain ATCC 25586."				
RL	J. Bacteriol. 184:2005-2018(2002).				
CC	-!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.				
CC	-!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.				
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).				
CC	-!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; AE010506; AAL94118.1; -.				
DR	HAMAP; MF_00328; -; 1.				
DR	InterPro; IPR000619; Guanylate_kin.				
DR	Pfam; PF00625; Guanylate_kin; 1.				
DR	SMART; SM00072; GuKC; 1.				
DR	PROSITE; PS00856; GUANYLATE_KINASE_1; 1.				
DR	PROSITE; PS50052; GUANYLATE_KINASE_2; 1.				
KW	Transferase; Kinase; ATP-binding; Complete proteome.				
FT	NP BIND 11 18 ATP (BY SIMILARITY).				
SQ	SEQUENCE 185 AA; 21025 MW; 988EF86443662B96 CRC64;				

Query Match 86.1%; Score 31; DB 1; Length 185;
Best Local Similarity 100.0%; Pred.No. 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 6
| | | | |
Db 131 RLGRNQ 136

```
RESULT 2
KGUA_ECO57
ID KGUA_ECO57 STANDARD; PRT; 207 AA.
AC Q8XD88;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMK OR Z5074 OR ECS4523.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RT Nature 409:529-533 (2001).
RL [2]
SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 050952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
CC -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE005592; AAG58792.1; -.
DR EMBL; AP002566; BAB37946.1; -.
DR PIR; C91194; C91194.
DR PIR; D86041; D86041.
DR HAMAP; MF_00328; -.
DR InterPro; IPR000619; Guanylate_kin.
DR Pfam; PF00625; Guanylate_kin; 1.
DR SMART; SM00072; GUKC; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 11 18 ATP (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23651 MW; CE669DB404E481E6 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 207;
Best Local Similarity 85.7%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
| | | | |
Db 134 RLGRDQ 140
```

```
RESULT 3
YH13_STRPN
ID YH13_STRPN STANDARD; PRT; 157 AA.
AC P58260;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein SP1713/spri1557.
GN SP1713 OR SPRI557.
OS Streptococcus pneumoniae, and
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313, 171101;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506 (2001).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-255 / R6;
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717 (2001).
CC -!- SIMILARITY: BELONGS TO THE UPF0168 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE007464; AAK75791.1; -.
DR EMBL; AE008523; AAL00361.1; ALT_INIT.
DR PIR; F95199; F95199.
DR TIGR; SP1713; -.
DR HAMAP; MF_00440; -.
DR InterPro; IPR005144; ATP.
DR InterPro; IPR003796; DUF193.
DR Pfam; PF03477; ATP-cone; 1.
DR TIGRFAMs; TIGR00244; TIGR00244; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 157 AA; 18380 MW; 7AF69B2A7E47249F CRC64;

Query Match 83.3%; Score 30; DB 1; Length 157;
Best Local Similarity 71.4%; Pred. No. 8.4;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
| | | | |
```


Db 94 KLRGRNE 100

RESULT 4

KGUA_ECOLI STANDARD; PRT; 207 AA.

AC P24234;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).

GN GMK OR SPOR OR B3648 OR C4473 OR SF3688.

OS Escherichia coli,

OS Escherichia coli O6, and

OS Shigella flexneri.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562, 217992, 623;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12;

RX MEDLINE=93300828; PubMed=8390989;

RA Gentry D., Bengra C., Ikehara K., Cashel M.;

RA "Guanylate kinase of Escherichia coli K-12.";

RL J. Biol. Chem. 268:14316-14321(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=K12 / MG1655;

RX MEDLINE=93315143; PubMed=7686882;

RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;

RA "DNA sequence and analysis of 136 kilobases of the Escherichia coli genome: organizational symmetry around the origin of replication.";

RL Genomics 16:551-561(1993).

RN [3]

RP SEQUENCE FROM N.A.

RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=22388234; PubMed=12471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RA "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).

RN [4]

RP SEQUENCE FROM N.A.

RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;

RX MEDLINE=2272406; PubMed=12384590;

RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,

RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,

RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,

RA Yu J.;

RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157.";

RL Nucleic Acids Res. 30:4432-4441(2002).

CC -!- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.

CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.

CC -!- SUBUNIT: HOMOTETRAMER (UNDER LOW IONIC CONDITIONS) OR HOMODIMER (UNDER HIGH IONIC CONDITIONS).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC -----

DR EMBL; M84400; AAB88711.1; -.

DR EMBL; L10328; AAA62001.1; -.

DR EMBL; AE000442; AAC76672.1; -.

DR EMBL; AE016769; AAN82909.1; ALT_INIT.

DR EMBL; AE015376; AAN45135.1; ALT_INIT.

DR PIR; S43041; KIECGU.

DR HSSP; P15454; 1GKY.

DR EcoGene; EG10965; gmk.

DR HAMAP; MF_00328; -; 1.

DR InterPro; IPR000619; Guanylate_kin.

DR Pfam; PF00625; Guanylate_kin; 1.

DR SMART; SM00072; GuKc; 1.

DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.

DR PROSITE; PS0052; GUANYLATE_KINASE_2; 1.

KW Transferase; Kinase; ATP-binding; Complete proteome.

FT NP_BIND 11 18 ATP (BY SIMILARITY).

SQ SEQUENCE 207 AA; 23592 MW; 62A99DB4063651E4 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 207;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 134 RLRGRGQ 140

RESULT 5

KGUA_SALTI STANDARD; PRT; 207 AA.

ID KGUA_SALTI

AC Q8Z2H9;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).

GN GMK OR STY4052 OR T3778.

OS Salmonella typhi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=601;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CT18;

RX MEDLINE=21534947; PubMed=11677608;

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,

RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,

RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,

RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,

RA Whitehead S., Barrell B.G.;

RA "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.";

RL Nature 413:848-852(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Ty2 / ATCC 700931;

RX MEDLINE=22531367; PubMed=12644504;

RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,

RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;

RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18.";

RL J. Bacteriol. 185:2330-2337(2003).

CC -!- FUNCTION: Essential for recycling GMP and indirectly, CGMP.

CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC or send an email to license@isb-sib.ch.

CC -----

DR EMBL; M84400; AAB88711.1; -.

DR EMBL; L10328; AAA62001.1; -.

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL627280; CAD03253.1; -.

DR EMBL; AE016846; AAO71261.1; -.

DR HAMAP; MF_00328; -; 1.

DR InterPro; IPR000619; Guanylate_kin.

DR Pfam; PF00625; Guanylate_kin; 1.

DR SMART; SM00072; GuKc; 1.

DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.

DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.

KW Transferase; Kinase; ATP-binding; Complete proteome.

FT NP_BIND 11 18 ATP (BY SIMILARITY).

SQ SEQUENCE 207 AA; 23499 MW; A26757C1FB40970B CRC64;

Query Match 83.3%; Score 30; DB 1; Length 207;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 134 RLRGRGQ 140

RESULT 6

KGUA_SALTY STANDARD; PRT; 207 AA.

AC Q9X6M5;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).

GN GMK OR STM3740.

OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Salmonella.

OX NCBI_TaxID=602;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2;

RA Beck B.J., Huelsmeyer M., Downs D.M.;

RA "Salmonella typhimurium guanylate kinase.";

RT Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=LT2 / SGSC1412 / ATCC 700720;

RC MEDLINE=21534948; PubMed=11677609;

RX McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;

RA "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";

RL Nature 413:852-856(2001).

CC -!- FUNCTION: ESSENTIAL FOR RECYCLING GMP AND INDIRECTLY, CGMP.

CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.

CC -!- SUBUNIT: HOMOTETRAMER (UNDER LOW IONIC CONDITIONS) OR HOMODIMER (UNDER HIGH IONIC CONDITIONS) (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF140283; AAD31506.1; -.

DR EMBL; AE008874; AAL22599.1; -.

DR HSSP; P15454; 1GKY.

DR StyGene; SG????; gmk.

DR HAMAP; MF_00328; -; 1.

DR InterPro; IPR000619; Guanylate_kin.

DR Pfam; PF00625; Guanylate_kin; 1.

DR SMART; SM00072; GuKc; 1.

DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.

DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.

KW Transferase; Kinase; ATP-binding; Complete proteome.

FT NP_BIND 11 18 ATP (BY SIMILARITY).

SQ SEQUENCE 207 AA; 23498 MW; A26FF5C1FB40970B CRC64;

Query Match 83.3%; Score 30; DB 1; Length 207;

Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 134 RLRGRGQ 140

RESULT 7

KGUA_YERPE STANDARD; PRT; 207 AA.

AC Q8ZJQ2;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).

GN GMK OR YPO0040 OR Y0101.

OS Yersinia pestis.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Yersinia.

OX NCBI_TaxID=632;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CO-92 / Biovar Orientalis;

RX MEDLINE=21470413; PubMed=11586360;

RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G., Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L., Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M., Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G., Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V., Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;

RT "Genome sequence of Yersinia pestis, the causative agent of plague.";

RL Nature 413:523-527(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=KIMS / Biovar Mediaevalis;

RX MEDLINE=22137863; PubMed=12142430;

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P., Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C., Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V., Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R., Perry R.D.;

RA "Genome sequence of Yersinia pestis KIM.";

RL J. Bacteriol. 184:4601-4611(2002).

CC -!- FUNCTION: Essential for recycling GMP and indirectly, CGMP.

CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO THE GUANYLATE KINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AJ414141; CAC8907.1; -.

DR EMBL; AE013610; AAM83695.1; -.

DR PIR; A10005; A10005.

DR HAMAP; MF_00328; -; 1.

DR InterPro; IPR000619; Guanylate kin.
DR Pfam; PF00625; Guanylate_kin; 1.
DR SMART; SM00072; GUKC; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 11 18 ATP (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23503 MW; 01FDCD618EFED91E CRC64;

Query Match 83.3%; Score 30; DB 1; Length 207;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 134 RLRGRGQ 140
|||||

RESULT 8
DNAB_MYCIT STANDARD; PRT; 576 AA.
AC Q9F5P4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Replicative DNA helicase (EC 3.6.1.-) [Contains: Min dnaB intein]
DE (Fragment).
DE DNAB.
GN Mycobacterium intracellulare.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1442;
RA Madiraju M.V.V.S., Yamamoto K., Rajagopalan M., Rutherford S.A.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PARTICIPATES IN INITIATION AND ELONGATION DURING
CC CHROMOSOME REPLICATION; IT EXHIBITS DNA-DEPENDENT ATPASE ACTIVITY
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HELICASE FAMILY. DNAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF307984; AAG31144.1; -.
DR InterPro; IPR001198; DnaB_helicase.
DR InterPro; IPR006141; Intein.
DR InterPro; IPR006142; INTEIN.
DR InterPro; IPR004042; Intein_endonuc.
DR Pfam; PF03796; DnaB_C; 1.
DR PRINTS; PR00379; INTEIN.
DR ProDom; PD32834; DnaB_C; 1.
DR TIGRFAMs; TIGR01443; Intein_Cterm; 1.
DR PROSITE; PS50818; INTEIN_C_TER; FALSE NEG.
DR PROSITE; PS50819; INTEIN_ENDONUCLEASE; 1.
DR PROSITE; PS50817; INTEIN_N_TER; FALSE NEG.
KW DNA replication; DNA-binding; ATP-binding; Hydrolase; Helicase;
KW Primosome; Autocatalytic cleavage; Protein splicing.
FT NON_TER 1 15
FT CHAIN <1 15 REPLICATIVE DNA HELICASE, 1ST PART
FT (POTENTIAL).
FT CHAIN 16 351 MIN DNAB INTEIN (POTENTIAL).
FT CHAIN 352 576 REPLICATIVE DNA HELICASE, 2ND PART
FT (POTENTIAL).

SQ SEQUENCE 576 AA; 63529 MW; 6D59341A779D082F CRC64;

Query Match 83.3%; Score 30; DB 1; Length 576;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 268 RVRGRNE 274
|:|||||

RESULT 9
UMPK_ARATH STANDARD; PRT; 202 AA.
ID UMPK_ARATH
AC O04905;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Uridylate kinase (EC 2.7.4.-) (UK) (Uridine monophosphate kinase)
DE (UMP kinase) (UMP/CMP kinase).
DE PYR6.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=98245083; PubMed=9576794;
RA Zhou L., Lacroute F., Thornburg R.;
RT "Cloning, expression in Escherichia coli, and characterization of
RT Arabidopsis thaliana UMP/CMP kinase.";
RL Plant Physiol. 117:245-254(1998).
CC -!- FUNCTION: FURNISHES THE CELL WITH PYRIMIDINES. ACCEPTS UMP AND
CC CMP AS PHOSPHATE ACCEPTORS WITH HIGH ACTIVITY; DOES NOT ACT ON
CC DUMP AND DUMP.
CC -!- CATALYTIC ACTIVITY: ATP + UMP = ADP + UDP.
CC -!- CATALYTIC ACTIVITY: ATP + CMP = ADP + CDP.
CC -!- SIMILARITY: Belongs to the adenylylate kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF000147; AAB71135.1; -.
DR HSSP; P20425; 3UKD.
DR InterPro; IPR000850; Adenylylate_kin.
DR InterPro; IPR006266; UMP_CMP_kinase.
DR Pfam; PF00406; ADK; 1.
DR PRINTS; PR00094; ADENYLTKINASE.
DR ProDom; PD000657; Adenylylate_kin; 1.
DR TIGRFAMs; TIGR01359; UMP_CMP_kin_fam; 1.
DR PROSITE; PS00113; ADENYLATE_KINASE; 1.
KW Transferase; Kinase; Pyrimidine biosynthesis; ATP-binding.
FT NP_BIND 21 29 ATP.
SQ SEQUENCE 202 AA; 22482 MW; 41AD0CFACD816315 CRC64;

Query Match 80.6%; Score 29; DB 1; Length 202;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 133 RLRGRNQ 139
|||||

RESULT 10
TF2B_ORYSA STANDARD; PRT; 312 AA.
ID TF2B_ORYSA

AC Q8W0W3; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription initiation factor IIB (General transcription factor
DE TFIIIB).
GN TFIIIB.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriophytidae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21967784; PubMed=11971135;
RA Zhu Q., Ordiz M.I., Dabi T., Beachy R.N., Lamb C.;
RT "Rice TATA binding protein interacts functionally with transcription
RT factor IIB and the RF2a bZIP transcriptional activator in an enhanced
RT plant in vitro transcription system."
RL Plant Cell 14:795-803 (2002).
CC -!- FUNCTION: GENERAL FACTOR THAT PLAYS A MAJOR ROLE IN THE ACTIVATION
CC OF EUKARYOTIC GENES TRANSCRIBED BY RNA POLYMERASE II.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: ASSOCIATES WITH TFIID-11A (DA COMPLEX) TO FORM TFIID-
CC 11A-11B (DAB-COMPLEX) WHICH IS THEN RECOGNIZED BY POLYMERASE II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the TFIIB family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF464908; AAL73491.1; --
DR Gramene; Q8W0W3; --
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR000812; TFIIB euk.
DR Pfam; PF00382; transcript_fac2; 2.
DR PRINTS; PR00685; TIFACTOR_IIB.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00782; TFIIB; 1.
KW Transcription regulation; Nuclear protein; Repeat; Zinc-finger;
KW Metal-binding; Zinc.
FT ZN FING 6 29 ZN-RIBBON TFIIB-TYPE.
FT REPEAT 115 192 1.
FT REPEAT 216 290 2.
FT METAL 6 6 ZINC (BY SIMILARITY).
FT METAL 9 9 ZINC (BY SIMILARITY).
FT METAL 26 26 ZINC (BY SIMILARITY).
FT METAL 29 29 ZINC (BY SIMILARITY).
SQ SEQUENCE 312 AA; 34298 MW; 338435FE048C46AF CRC64;
Query Match 80.6%; Score 29; DB 1; Length 312;
Best Local Similarity 83.3%; Pred. No. 30;
Matches .5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 2 LRGRNQ 7
Db 142 LRGRNQ 147
RESULT 11
LEGA_PEA STANDARD; PRT; 517 AA.
AC P02857; Q41033; Q41034;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Legumin A precursor.
GN LEGA.

OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Feltham First;
RX MEDLINE=84247316; PubMed=6330672;
RA Lyett G.W., Croy R.R.D., Shirsat A.H., Boulter D.;
RT "The complete nucleotide sequence of a legumin gene from pea (Pisum
RT sativum L.).";
RL Nucleic Acids Res. 12:4493-4506 (1984).
RN [2]
RP SEQUENCE OF 209-411 FROM N.A.
RA Domoney C., Barker D., Casey R.;
RT "The complete deduced amino acid sequences of legumin beta-
RT polypeptides from different genetic loci in Pisum.";
RL Plant Mol. Biol. 7:467-474 (1986).
CC -!- FUNCTION: THIS PROTEIN FOUND IN THE SEEDS OF MANY LEGUMINOUS &
CC NONLEGUMINOUS PLANTS IS THE SOURCE OF SULFUR-CONTAINING AMINO
CC ACIDS IN SEED MEALS.
CC -!- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
CC DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the IIS seed storage protein (globulins)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X02982; CAA26720.1; ALT_SEQ.
DR EMBL; M16904; AAA33677.1; --
DR PIR; A22866; FWPMLA.
DR PIR; S09559; S09559.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_sup.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; IIS SEED STORAGE; 1.
KW Seed storage protein; Multigene family; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 332 ALPHA CHAIN (ACIDIC).
FT CHAIN 333 517 BETA CHAIN (BASIC).
FT DISULFID 107 339 INTERCHAIN (ALPHA-BETA) (POTENTIAL).
FT DOMAIN 89 122 GLN/GLU-RICH.
FT DOMAIN 189 208 GLN-RICH.
FT DOMAIN 260 313 ARG/GLN/GLU-RICH.
FT CONFLICT 219 219 Y -> F (IN REF. 2).
FT CONFLICT 319 319 G -> R (IN REF. 2).
SQ SEQUENCE 517 AA; 58805 MW; CD2CA2B41758D0BA CRC64;
Query Match 80.6%; Score 29; DB 1; Length 517;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RLGRNQ 7
Db 233 RLGRNE 239
RESULT 12
LEGA_PEA STANDARD; PRT; 520 AA.
ID LEG2_PEA
AC P15838;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Legumin A2 precursor.
GN LEGA2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eursids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Greenfeast;
RX MEDLINE=90174993; PubMed=2308850;
RA Rerie W.G.; Whitecross M.I.; Higgins T.J.V.;
RT "Nucleotide sequence of an A-type legumin gene from pea."
RL Nucleic Acids Res. 18:655-655(1990).
CC -!- FUNCTION: THIS PROTEIN FOUND IN THE SEEDS OF MANY LEGUMINOUS &
CC NONLEGUMINOUS PLANTS IS THE SOURCE OF SULFUR-CONTAINING AMINO
CC ACIDS IN SEED MEALS.
CC -!- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
CC BASIC CHAIN DERIVED FROM A SINGLE PRECURSOR AND LINKED BY A
CC DISULFIDE BOND.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17193; CAA35056.1; --
DR PIR; S08237; S08237.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_sup.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PRO0439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW Seed storage protein; Multigene family; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 335 ALPHA CHAIN (ACIDIC).
FT CHAIN 336 520 BETA CHAIN (BASIC).
FT DISULFID 108 342 INTERCHAIN (ALPHA-BETA) (POTENTIAL).
SQ SEQUENCE 520 AA; 59269 MW; 59269 MW; CDB8B5E350C760C6 CRC64;
Query Match 80.6%; Score 29; DB 1; Length 520;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RLGRNQ 7
Db 234 RLQGRNE 240
RESULT 13
COAA BACTJ
ID COAA BACTJ STANDARD; PRT; 674 AA.
AC O87905;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pesticidal crystal protein cry24Aa (Insecticidal delta-endotoxin
DE CryXXIVA(a)) (Crystalline entomocidal protoxin) (Crystal protein)
DE (Insecticidal protein Jegg72) (Fragment).
GN CRY24AA OR CRYXXIVA(A).
OS Bacillus thuringiensis (subsp. jegathesan).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=56955;
RN [1]
RP SEQUENCE FROM N.A.
RX Kawalek M.D.; Gill S.S.;

RT "Isolation and characterization of insecticidal genes from Bacillus
thuringiensis subsp. jegathesan."
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -!- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -!- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -!- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U88188; AAC61891.1; --
DR InterPro; IPR001178; Endotoxin.
DR InterPro; IPR005638; endotoxin_N.
DR InterPro; IPR005639; endotoxin_N.
DR Pfam; PF00555; endotoxin; 1.
DR Pfam; PF03944; endotoxin_C; 1.
DR Pfam; PF03945; endotoxin_N; 1.
KW Toxin; Sporulation.
FT NON_TER 674 674
SQ SEQUENCE 674 AA; 75959 MW; DA3904DAB891C978 CRC64;
Query Match 80.6%; Score 29; DB 1; Length 674;
Best Local Similarity 83.3%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2 LRGRNQ 7
Db 435 LRGRNQ 440
RESULT 14
CV02_HUMAN
ID CV02_HUMAN STANDARD; PRT; 126 AA.
AC Q9Y3M2; Q9UIK9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein C22orf2 (Cytosolic leucine-rich protein) (HRIHFB2025).
GN C22ORF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S.; Weil B.; Wellenreuther R.; Gassenhuber J.; Glassl S.;
RA Ansorge W.; Boecker M.; Bloecker H.; Bauersachs S.; Blum H.;
RA Lauber J.; Duesterhoeft A.; Beyer A.; Koehrer K.; Strack N.;
RA Mewes H.-W.; Ottenwaelder B.; Obermaier B.; Tampe J.; Heubner D.;
RA Wambutt R.; Korn B.; Klein M.; Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs."
RL Genome Res. 11:422-435(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Collins J.E.; Huckle E.J.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
Dunham I.; Hunt A.R.; Collins J.E.; Bruskewich R.; Beare D.M.;

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- TISSUE SPECIFICITY: Found in heart, brain, lung, liver, muscle,
CC kidney and testis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF331040; AAL56061.1; -.
DR EMBL; AK003719; BAB22956.1; -.
DR EMBL; BC005733; AAH05733.1; -.
FT DOMAIN 77 98 LEUCINE-ZIPPER.
SQ SEQUENCE 127 AA; 14534 MW; 7CE38A48D045CCC7 CRC64;

Query Match 77.8%; Score 28; DB 1; Length 127;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
||| |||
Db 76 RLRRNQ 82

Search completed: February 11, 2004, 17:04:12
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLRGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	31	86.1	75	2 AB1193	hypothetical prote
2	31	86.1	145	2 AD2345	hypothetical prote
3	31	86.1	195	2 F96734	probable transcrip
4	31	86.1	207	2 C91194	guanylate kinase
5	31	86.1	207	2 D86041	guanylate kinase
6	31	86.1	213	1 S71287	myb-related 24.7K
7	31	86.1	280	2 AF1375	B. subtilis YunF p
8	31	86.1	281	2 AD1745	B. subtilis YunF p
9	31	86.1	386	2 AE1996	hypothetical prote
10	30	83.3	157	2 F95199	conserved hypothet
11	30	83.3	177	2 D98066	conserved hypothet
12	30	83.3	207	1 KIECGU	guanylate kinase
13	30	83.3	207	2 AI0005	guanylate kinase
14	30	83.3	207	2 AI0969	5'guanylate kinase
15	30	83.3	445	2 T21262	hypothetical prote
16	30	83.3	1050	2 T26395	hypothetical prote
17	29	80.6	170	2 AB2402	hypothetical prote
18	29	80.6	184	2 B42707	transposase homolo
19	29	80.6	203	2 S09559	legumin A (clone p
20	29	80.6	203	2 F72273	conserved hypothet
21	29	80.6	251	2 AB0213	conserved hypothet
22	29	80.6	307	2 PC2050	legumin type A alp
23	29	80.6	315	2 JC2094	legumin type A alp
24	29	80.6	482	2 S49877	legumin A precursor
25	29	80.6	497	2 S14392	legumin A1 precursor
26	29	80.6	498	2 S44294	legumin A precursor
27	29	80.6	500	2 S14393	legumin A2 precursor
28	29	80.6	507	2 T06452	probable legumin A
29	29	80.6	517	1 FWPVLA	legumin A precursor

30	29	80.6	520	2 S08237	legumin A2 precurs
31	29	80.6	779	2 S36111	osteoblast-specifi
32	29	80.6	811	2 S36109	osteoblast-specifi
33	29	80.6	836	2 S36110	osteoblast-specifi
34	29	80.6	2610	2 T20968	hypothetical prote
35	28	77.8	75	2 AH1550	hypothetical prote
36	28	77.8	123	2 D84161	hypothetical prote
37	28	77.8	144	2 T49706	hypothetical prote
38	28	77.8	178	2 H69499	conserved hypothet
39	28	77.8	193	2 E83864	thioredoxin BH1717
40	28	77.8	199	2 AB1820	guanylate kinase
41	28	77.8	242	2 AB1866	tRNA nucleotidyltr
42	28	77.8	242	2 G81076	ribonuclease PH NM
43	28	77.8	260	1 A38585	deoxycytidine kina
44	28	77.8	260	2 A55122	deoxycytidine kina
45	28	77.8	278	2 AD2325	hypothetical prote

ALIGNMENTS

RESULT 1

AB1193
hypothetical protein lmo0946 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AB1193
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, F
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlanc
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1193
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-75 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99024.1; PID:g16410348; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo0946

Query Match 86.1%; Score 31; DB 2; Length 75;
Best Local Similarity 100.0%; Pred.No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
|||||
Db 70 RLRGRN 75

RESULT 2

AD2345
hypothetical protein all4315 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD2345
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2345
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-145 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA076014.1; PID:g17133451; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4315.

Db 134 RLRGRDQ 140

RESULT 5

D86041

guanylate kinase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C;Accession: D86041

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: D86041

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <STO>

A;Cross-references: GB:AE005174; NID:gl2518407; PIDN:AAG58792.1; GSPDB:GN00145; UWGP:Z5

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: gmk

C;Superfamily: guanylate kinase; guanylate kinase homology

Query Match 86.1%; Score 31; DB 2; Length 207;

Best Local Similarity 85.7%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 134 RLRGRDQ 140

RESULT 6

S71287

myb-related 24.7K transcription factor - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 16-Jun-2000

C;Accession: S71287; JC5729

R;Kirik, V.; Baumlein, H.

submitted to the EMBL Data Library, December 1995

A;Description: Isolation and chromosomal localization of leaf-specific cDNA encoding fc

A;Reference number: S71287

A;Accession: S71287

A;Molecule type: mRNA

A;Residues: 1-213 <KIR>

A;Cross-references: EMBL:Z68157; NID:g1197189; PIDN:CAA92280.1; PID:g1197190

R;Kirik, V.; Baeumlein, H.

Gene 183, 109-113, 1996

A;Title: A novel leaf-specific myb-related protein with a single binding repeat.

A;Reference number: JC5729; MUID:97149286; PMID:8996094

A;Accession: JC5729

A;Molecule type: mRNA

A;Residues: 1-213 <K12>

A;Cross-references: EMBL:Z68157; NID:g1197189; PIDN:CAA92280.1; PID:g1197190

C;Comment: This protein is involved in the control of anthocyanin biosynthesis and the

C;Genetics:

A;Gene: AtmybL2

A;Map position: 1

C;Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology

C;Keywords: DNA binding; nucleus; transcription regulation

F;30-80/Domain: myb DNA-binding repeat homology <MYB2>

F;54-80/Region: helix-turn-helix

Query Match 86.1%; Score 31; DB 1; Length 213;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 22 RLRGRN 27

Query Match 86.1%; Score 31; DB 2; Length 207;

Best Local Similarity 85.7%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 22 RLRGRN 27

Query Match 86.1%; Score 31; DB 2; Length 145;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 99 RLRGRN 104

RESULT 3

F96734

probable transcription factor F23N20.2 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001

C;Accession: F96734

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Mafti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: F96734

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-195 <STO>

A;Cross-references: GB:AE005173; NID:g6714312; PIDN:AAF26005.1; GSPDB:GN00141

C;Genetics:

A;Gene: F23N20.2

A;Map position: 1

C;Superfamily: Arabidopsis myb-related 24.7K protein; myb DNA-binding repeat homology

Query Match 86.1%; Score 31; DB 2; Length 195;

Best Local Similarity 100.0%; Pred. No. 23;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 22 RLRGRN 27

RESULT 4

C91194

guanylate kinase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C;Accession: C91194

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: C91194

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA037946.1; PID:g133639998; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs4523

C;Superfamily: guanylate kinase; guanylate kinase homology

Query Match 86.1%; Score 31; DB 2; Length 207;

Best Local Similarity 85.7%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 22 RLRGRN 27

RESULT 7

AF1375
B. subtilis YunF protein homolog lmo2406 [imported] - Listeria monocytogenes (strain EGD)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AF1375
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1375
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00484.1; PID:gl6411894; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2406

Query Match 86.1%; Score 31; DB 2; Length 280;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||
Db 192 RLHGRNQ 198

RESULT 8

AD1745
B. subtilis YunF protein homolog lin2505 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AD1745
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97732.1; PID:gl6415027; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2505

Query Match 86.1%; Score 31; DB 2; Length 281;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|||
Db 193 RLHGRNQ 199

RESULT 9

AE1996
hypothetical protein all1523 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE1996
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ar
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1996
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-386 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB77889.1; PID:gl7135343; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1523

Query Match 86.1%; Score 31; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRN 6
|||
Db 299 RLGRN 304

RESULT 10

F95199
conserved hypothetical protein SP1713 [imported] - Streptococcus pneumoniae (strain TIC)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C;Accession: F95199
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrisor
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: F95199
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <KUR>
A;Cross-references: GB:AF005672; PIDN:AAK75791.1; PID:gl4973208; GSPDB:GN00164; TIGR:SE
A;Experimental source: strain TIGR
C;Genetics:
A;Gene: SP1713
C;Superfamily: conserved hypothetical protein HI0943

Query Match 83.3%; Score 30; DB 2; Length 157;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
:|:|:|:
Db 94 KLRGRNE 100

RESULT 11

D98066
conserved hypothetical protein spr1557 [imported] - Streptococcus pneumoniae (strain R6
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: D98066
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;
e, R.; LeBlanc, D.J.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D98066
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-177 <KUR>
A;Cross-references: GB:AB007317; PIDN:AAL00361.1; PID:gl5459222; GSPDB:GN00174
C;Genetics:
A;Gene: spr1557

C;Superfamily: conserved hypothetical protein HI0943

Query Match 83.3%; Score 30; DB 2; Length 177;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
:|||||
Db 114 KLRGRNE 120

RESULT 12

KIECGU

Guanylate kinase (EC 2.7.4.8) - Escherichia coli (strain K-12)

C;Species: Escherichia coli
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 01-Mar-2002
C;Accession: S43041; B65166; S24192
R;Gentry, D.; Bengra, C.; Ikehara, K.; Cashel, M.
J. Biol. Chem. 268, 14316-14321, 1993

A;Title: Guanylate kinase of Escherichia coli K-12.

A;Reference number: S43040; MUID:93300828; PMID:8390989

A;Accession: S43041

A;Molecule type: DNA

A;Residues: 1-207 <GEN>

A;Cross-references: EMBL:M84400; NID:gl46228; PIDN:AAB89711.1; PID:gl46230

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B65166

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-207 <BLAT>

A;Cross-references: GB:AE000442; GB:U00096; NID:g2367253; PIDN:AAC76672.1; PID:gl790080;

A;Experimental source: strain K-12, substrain MGI655

C;Genetics:

A;Gene: gmk

C;Superfamily: guanylate kinase; guanylate kinase homology

C;Keywords: ATP; magnesium; monomer; nucleotide binding; P-loop; phosphotransferase

F;5-187/Domain: guanylate kinase homology <GKI>

F;11-18/Region: nucleotide-binding motif A (P-loop)

F;37-84/Region: GMP binding #status predicted

F;17/Binding site: ATP (Lys) #status predicted

Query Match 83.3%; Score 30; DB 1; Length 207;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
:|||||
Db 134 RLRGRGQ 140

RESULT 13

AI0005

Guanylate kinase (EC 2.7.4.8) [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001

C;Accession: AI0005

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AI0005

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC88907.1; PID:gl5978155; GSPDB:GN00175

C;Genetics:

A;Gene: gmk

C;Superfamily: guanylate kinase; guanylate kinase homology
C;Keywords: phosphotransferase

Query Match 83.3%; Score 30; DB 2; Length 207;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
:|||||
Db 134 RLRGRGQ 140

RESULT 14

AI0969

5'guanylate kinase [imported] - Salmonella enterica subsp. enterica serovar Typhi (stra

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AI0969

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher

th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar

, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serc

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AI0969

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD03253.1; PID:gl6504875; GSPDB:GN00176

C;Genetics:

A;Gene: STY4052

C;Superfamily: guanylate kinase; guanylate kinase homology

Query Match 83.3%; Score 30; DB 2; Length 207;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
:|||||
Db 134 RLRGRGQ 140

RESULT 15

T21262

hypothetical protein F22D6.11 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T21262

R;Wilkinson, J.

submitted to the EMBL Data Library, April 1996

A;Reference number: Z19397

A;Accession: T21262

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-445 <WIL>

A;Cross-references: EMBL:T21262; PIDN:CAA95817.1; GSPDB:GN00019; CESP:F22D6.11

A;Experimental source: clone F22D6

C;Genetics:

A;Gene: CESP:F22D6.11

A;Map position: 1

A;introns: 65/3; 95/2; 136/3; 180/3; 256/3; 281/3; 403/3

Query Match 83.3%; Score 30; DB 2; Length 445;
Best Local Similarity 71.4%; Pred. No. 87;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
:|||||
Db 251 RLRGKNE 257

Search completed: February 11, 2004, 17:11:42
Job time : 10.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap:
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pap:
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pap:
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap:
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pap:
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap:
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pap:
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap:
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap:
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap:
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap:
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pap:
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap:
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap:
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap:
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pap:
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pap:
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	7	10	US-09-901-187B-8
2	33	91.7	84	12	US-09-864-408A-8020
3	29	80.6	129	15	US-10-106-698-4874
4	29	80.6	696	15	US-10-176-847-46
5	29	80.6	730	15	US-10-217-371-6
6	29	80.6	750	15	US-10-217-371-14
7	29	80.6	751	15	US-10-217-371-4
8	29	80.6	758	15	US-10-217-371-10
9	29	80.6	771	15	US-10-217-371-12
10	29	80.6	779	12	US-10-301-822-151
11	29	80.6	779	15	US-10-217-371-8
12	29	80.6	779	15	US-10-171-311-180
13	29	80.6	790	9	US-09-925-301-1313
14	29	80.6	836	12	US-10-301-822-149
15	29	80.6	836	12	US-10-204-752-33

16	29	80.6	836	12	US-10-295-027-36	Sequence 36, Appl
17	29	80.6	836	12	US-10-173-999-66	Sequence 66, Appl
18	29	80.6	836	15	US-10-217-371-2	Sequence 2, Appl
19	29	80.6	836	15	US-10-171-311-178	Sequence 178, Appl
20	29	80.6	836	15	US-10-177-293-342	Sequence 342, Appl
21	29	80.6	2610	12	US-10-369-493-6553	Sequence 6553, Ap
22	29	80.6	2610	12	US-10-369-493-6554	Sequence 6554, Ap
23	28	77.8	54	11	US-09-832-355A-114	Sequence 114, Appl
24	28	77.8	168	12	US-10-094-749-2002	Sequence 2002, Ap
25	28	77.8	169	12	US-10-114-153-72	Sequence 72, Appl
26	28	77.8	199	12	US-10-136-728-52	Sequence 52, Appl
27	28	77.8	199	15	US-10-156-761-8025	Sequence 8025, Ap
28	28	77.8	289	9	US-09-925-301-897	Sequence 897, Appl
29	28	77.8	313	12	US-10-017-161-752	Sequence 752, Appl
30	28	77.8	313	12	US-10-041-615-74	Sequence 74, Appl
31	28	77.8	313	12	US-10-041-615-81	Sequence 81, Appl
32	28	77.8	316	12	US-10-292-798-656	Sequence 656, Appl
33	28	77.8	317	10	US-09-886-055-297	Sequence 297, Appl
34	28	77.8	317	11	US-09-804-291-297	Sequence 297, Appl
35	28	77.8	469	12	US-10-108-260A-2918	Sequence 2918, Ap
36	28	77.8	474	12	US-10-369-493-2299	Sequence 2299, Ap
37	28	77.8	503	15	US-10-225-060-11	Sequence 11, Appl
38	28	77.8	506	15	US-10-156-761-8426	Sequence 8426, Ap
39	28	77.8	509	15	US-10-215-224-2	Sequence 2, Appl
40	28	77.8	509	15	US-10-215-224-3	Sequence 3, Appl
41	28	77.8	509	15	US-10-225-060-10	Sequence 10, Appl
42	28	77.8	1038	12	US-10-369-493-1708	Sequence 1708, Ap
43	28	77.8	1134	14	US-10-001-873-50	Sequence 50, Appl
44	28	77.8	1253	12	US-10-369-493-20694	Sequence 20694, A
45	28	77.8	1257	10	US-09-738-626-4750	Sequence 4750, Ap

ALIGNMENTS

RESULT 1

US-09-901-187B-8

; Sequence 8, Application US/09901187B

; Patent No. US20020151464A1

; GENERAL INFORMATION:

; APPLICANT: Panacea Pharmaceuticals, Inc.

; APPLICANT: Wolozin, Benjamin

; APPLICANT: Ostretova-Golts, Natalie

; APPLICANT: Lebowitz, Micheal S.

; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment

; TITLE OF INVENTION: Alpha-Synuclein Diseases

; FILE REFERENCE: PAN01/002US

; CURRENT APPLICATION NUMBER: US/09/901,187B

; CURRENT FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/217,319

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/279,199

; PRIOR FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-901-187B-8

Query Match 100.0%; Score 36; DB 10; Length 7;

Best Local Similarity 100.0%; Pred. No. 7e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLGRNQ 7

|||||

Db 1 RLGRNQ 7

RESULT 2

US-09-864-408A-8020

; Sequence 8020, Application US/09864408A

```

; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8020
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-8020

Query Match          91.7%; Score 33; DB 12; Length 84;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLRGRNQ 7
       |:|||||
Db      31 RVRGRNQ 37

RESULT 3
US-10-106-698-4874
; Sequence 4874, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4874
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4874

Query Match          80.6%; Score 29; DB 15; Length 129;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLRGRNQ 7
       |:|||||
Db      4 RVRGRNR 10

RESULT 4
US-10-176-847-46
; Sequence 46, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-176-847-46

Query Match          80.6%; Score 29; DB 15; Length 696;
Best Local Similarity 71.4%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLRGRNQ 7
       |:|||||
Db      33 RIRGRDQ 39

RESULT 5
US-10-217-371-6
; Sequence 6, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-6

Query Match          80.6%; Score 29; DB 15; Length 730;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      1 RLRGRNQ 7
       |:|||||
Db      12 RIRGRDQ 18

RESULT 6
US-10-217-371-14
; Sequence 14, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-14

Query Match          80.6%; Score 29; DB 15; Length 750;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 12 RIRGRDQ 18

RESULT 7
US-10-217-371-4
; Sequence 4, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ian Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-4

Query Match 80.6%; Score 29; DB 15; Length 751;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 33 RIRGRDQ 39

RESULT 8
US-10-217-371-10
; Sequence 10, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ian Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-10

Query Match 80.6%; Score 29; DB 15; Length 758;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 12 RIRGRDQ 18

RESULT 9
US-10-217-371-12

; Sequence 12, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ian Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-12

Query Match 80.6%; Score 29; DB 15; Length 771;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 33 RIRGRDQ 39

RESULT 10
US-10-301-822-151
; Sequence 151, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burt, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-301-822-151

Query Match 80.6%; Score 29; DB 12; Length 779;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 33 RIRGRDQ 39

RESULT 11

US-10-217-371-8
; Sequence 8, Application US/10217371
; Publication No. US20030073137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Lan Bo
; APPLICANT: Dai, Meiru
; APPLICANT: Sasaki, Hidefumi
; APPLICANT: Auclair, Daniel
; TITLE OF INVENTION: PERIOSTIN-BASED DIAGNOSTIC ASSAYS
; FILE REFERENCE: 00530-099001
; CURRENT APPLICATION NUMBER: US/10/217,371
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US 60/312,123
; PRIOR FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-217-371-8

Query Match 80.6%; Score 29; DB 15; Length 779;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 12
US-10-171-311-180
; Sequence 180, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 180
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-180

Query Match 80.6%; Score 29; DB 15; Length 779;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 13
US-09-925-301-1313
; Sequence 1313, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1313
; LENGTH: 790
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1313

Query Match 80.6%; Score 29; DB 9; Length 790;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:|:|:|:
Db 44 RIRGRDQ 50

RESULT 14
US-10-301-822-149
; Sequence 149, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-149

Query Match 80.6%; Score 29; DB 12; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:|:|:|:
Db 33 RIRGRDQ 39

RESULT 15
US-10-204-752-33
; Sequence 33, Application US/10204752
; Publication No. US20030152956A1
; GENERAL INFORMATION:
; APPLICANT: OHTANI, No. US20030152956A1iko
; APPLICANT: MATSUI, Keiko
; APPLICANT: YOSHIDA, Nei
; APPLICANT: SUGITA, Yuji
; APPLICANT: IZUHARA, Kenji
; TITLE OF INVENTION: METHOD OF TESTING FOR ALLERGIC DISEASE
; FILE REFERENCE: SHZ-009US
; CURRENT APPLICATION NUMBER: US/10/204,752
; CURRENT FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: JP 2000-396166
; PRIOR FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-204-752-33

Query Match 80.6%; Score 29; DB 12; Length 836;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLRGRNQ 7
Db 33 RIRGRDQ 39

Search completed: February 11, 2004, 17:54:11
Job time : 25.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-8
Perfect score: 36
Sequence: 1 RLRGRNQ 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCITUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	88.9	494	4	US-09-252-991A-17295
2	30	83.3	195	4	US-09-252-991A-16682
3	30	83.3	281	4	US-09-252-991A-33048
4	30	83.3	357	4	US-09-252-991A-31718
5	30	83.3	435	4	US-09-252-991A-30142
6	30	83.3	528	4	US-09-252-991A-28291
7	29	80.6	779	1	US-08-426-627-4
8	29	80.6	779	1	US-08-426-627-24
9	29	80.6	811	1	US-08-426-627-2
10	29	80.6	811	1	US-08-426-627-22
11	29	80.6	836	1	US-08-426-627-6
12	29	80.6	837	1	US-08-426-627-23
13	28	77.8	260	2	US-08-879-561-8
14	28	77.8	262	4	US-09-857-498A-6
15	28	77.8	310	4	US-09-107-532A-4212
16	28	77.8	442	4	US-09-252-991A-31848
17	28	77.8	503	3	US-08-740-233A-11
18	28	77.8	503	4	US-09-709-188-11
19	28	77.8	509	2	US-08-665-926-8
20	28	77.8	509	3	US-08-740-223A-10
21	28	77.8	509	4	US-09-202-491-2
22	28	77.8	509	4	US-09-202-491-3
23	28	77.8	509	4	US-09-709-188-10
24	28	77.8	736	3	US-09-102-528-29
25	28	77.8	1038	3	US-09-541-782-4
26	28	77.8	1038	4	US-09-723-820-4
27	28	77.8	1184	3	US-09-541-782-2

28	77.8	1184	4	US-09-723-820-2	Sequence 2, Appli
29	77.8	1257	2	US-08-750-152A-2	Sequence 2, Appli
30	75.0	43	1	US-08-098-005-1	Sequence 1, Appli
31	75.0	73	1	US-07-609-716-109	Sequence 109, App
32	75.0	73	1	US-07-609-716-110	Sequence 110, App
33	75.0	73	3	US-08-475-411A-109	Sequence 109, App
34	75.0	73	3	US-08-475-411A-110	Sequence 110, App
35	75.0	73	3	US-08-478-029A-109	Sequence 110, App
36	75.0	73	3	US-08-478-029A-110	Sequence 110, App
37	75.0	97	1	US-07-609-716-99	Sequence 99, Appli
38	75.0	97	3	US-08-475-411A-99	Sequence 99, Appli
39	75.0	97	3	US-08-478-029A-99	Sequence 99, Appli
40	75.0	99	2	US-08-292-968-25	Sequence 25, Appli
41	75.0	99	2	US-08-467-974-25	Sequence 25, Appli
42	75.0	99	2	US-08-467-536-25	Sequence 25, Appli
43	75.0	99	3	US-08-467-976-25	Sequence 25, Appli
44	75.0	99	3	US-09-082-514-25	Sequence 25, Appli
45	75.0	100	2	US-08-849-376-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-17295
; Sequence 17295, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17295
; LENGTH: 494
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17295

Query Match 88.9%; Score 32; DB 4; Length 494;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLRGRNQ 7
Db 243 RLRGRNR 249

RESULT 2
US-09-252-991A-16682
; Sequence 16682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16682
; LENGTH: 195
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16682

Query Match 83.3%; Score 30; DB 4; Length 195;
Best Local Similarity 85.7%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
Db 21 RLRGRRQ 27

RESULT 3

US-09-252-991A-33048
; Sequence 33048, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 33048
; LENGTH: 281
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-33048

Query Match 83.3%; Score 30; DB 4; Length 281;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
Db 59 RLRGRRQ 65

RESULT 4

US-09-252-991A-31718
; Sequence 31718, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31718
; LENGTH: 357
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31718

Query Match 83.3%; Score 30; DB 4; Length 357;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
Db 48 RLRGRRQ 54

RESULT 5

US-09-252-991A-30142
; Sequence 30142, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30142
; LENGTH: 435
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30142

Query Match 83.3%; Score 30; DB 4; Length 435;
Best Local Similarity 85.7%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
Db 254 RLRGRGQ 260

RESULT 6

US-09-252-991A-28291
; Sequence 28291, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28291
; LENGTH: 528
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28291

Query Match 83.3%; Score 30; DB 4; Length 528;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RLRGRNQ 7
Db 170 RLRGRRQ 176

RESULT 7

US-08-426-627-4
; Sequence 4, Application US/08426627
; Patent No. 5756664

; GENERAL INFORMATION:

; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi

```
;
; TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,627
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
; REFERENCE/DOCKET NUMBER: 02481-1285-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-426-627-4

Query Match 80.6%; Score 29; DB 1; Length 779;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 33 RIRGRDQ 39

RESULT 8
US-08-426-627-24
; Sequence 24, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,627
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
; REFERENCE/DOCKET NUMBER: 02481-1285-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-426-627-24

Query Match 80.6%; Score 29; DB 1; Length 779;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 33 RIRGRDQ 39

RESULT 9
US-08-426-627-2
; Sequence 2, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,627
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
```

REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-426-627-2

Query Match 80.6%; Score 29; DB 1; Length 811;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 35 RIRGRDQ 41

RESULT 10
US-08-426-627-22
; Sequence 22, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,627
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
; REFERENCE/DOCKET NUMBER: 02481-1285-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 811 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-426-627-22

Query Match 80.6%; Score 29; DB 1; Length 811;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 RLRGRNQ 7
|:||||:
Db 35 RIRGRDQ 41

RESULT 11
US-08-426-627-6
; Sequence 6, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko
; APPLICANT: Kikuno, Reiko
; APPLICANT: Takeshita, Sunao
; APPLICANT: Tezuka, Kenichi
; TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
; TITLE OF INVENTION: Ability and Process for Its Production.
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,627
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,841
; FILING DATE: 25-MAR-1993
; APPLICATION NUMBER: JP 4-71501
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hammond, Alan W.
; REGISTRATION NUMBER: 35,178
; REFERENCE/DOCKET NUMBER: 02481-1285-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 836 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-426-627-6

Query Match 80.6%; Score 29; DB 1; Length 836;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|:||||:
Db 33 RIRGRDQ 39

RESULT 12
US-08-426-627-23
; Sequence 23, Application US/08426627
; Patent No. 5756664
; GENERAL INFORMATION:
; APPLICANT: Amann, Egon
; APPLICANT: Otawara-Hamamoto, Yoko

APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
APPLICANT: Tezuka, Kenichi
TITLE OF INVENTION: No. 5756664el Protein with Bone Formation
TITLE OF INVENTION: Ability and Process for Its Production.
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-426-627-23

Query Match 80.6%; Score 29; DB 1; Length 837;
Best Local Similarity 71.4%; Pred. No. 3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
|:|:|:
Db 33 RIRGRDQ 39

RESULT 13
US-08-879-561-8
; Sequence 8, Application US/08879561
; Patent No. 5817482
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 260 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 181510
US-08-879-561-8

Query Match 77.8%; Score 28; DB 2; Length 260;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRGRNQ 7
|:|:|:
Db 191 LRGRNE 196

RESULT 14
US-09-857-498A-6
; Sequence 6, Application US/09857498A
; Patent No. 6582947
; GENERAL INFORMATION:
; APPLICANT: Karlsson, Anna
; APPLICANT: Johansson, Magnus
; TITLE OF INVENTION: New medical use
; FILE REFERENCE: multisubstrate deoxyguanosine kinase
; CURRENT APPLICATION NUMBER: US/09/857,498A
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: SE 9804298-9
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-857-498A-6

Query Match 77.8%; Score 28; DB 4; Length 262;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRGRNQ 7
|:|:|:
Db 192 LRGRNE 197

RESULT 15
US-09-107-532A-4212
; Sequence 4212, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 4212:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 310 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...310
 SEQUENCE DESCRIPTION: SEQ ID NO: 4212:
 US-09-107-532A-4212

Query Match 77.8%; Score 28; DB 4; Length 310;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
 |||||
 Db 277 RLGENE 283

Search completed: February 11, 2004, 17:13:37
 Job time: 10.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-8

Perfect score: 36

Sequence: 1 RLRGRNQ 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	7	AAE14553	Human alpha-synuclein
2	36	100.0	533	ABP63409	Drosophila melanog
3	33	91.7	84	ABP35037	Human ORF4010 prot
4	32	88.9	458	ABG17799	Novel human diagno
5	31	86.1	75	ABP49607	Listeria monocytog
6	31	86.1	78	AAU61300	Propionibacterium
7	31	86.1	87	ABG20695	Novel human diagno
8	31	86.1	131	AAG18501	Zea mays protein f
9	31	86.1	139	AAG25683	Arabidopsis thalia

10	31	86.1	145	21	AAG18500	Zea mays protein f
11	31	86.1	153	21	AAG25682	Arabidopsis thalia
12	31	86.1	181	21	AAG16627	Arabidopsis thalia
13	31	86.1	195	21	AAG16626	Arabidopsis thalia
14	31	86.1	280	23	ABB50063	Listeria monocytog
15	31	86.1	337	22	AAU46599	Propionibacterium
16	31	86.1	369	22	AAU62018	Propionibacterium
17	31	86.1	384	22	ABB67872	Drosophila melanog
18	31	86.1	472	22	ABB66583	Drosophila melanog
19	31	86.1	493	21	AAU13570	Streptomyces globi
20	30	83.3	67	19	AAW80711	S. pneumoniae prot
21	30	83.3	123	21	AAG19228	Arabidopsis thalia
22	30	83.3	131	21	AAG19227	Arabidopsis thalia
23	30	83.3	150	21	AAG19226	Arabidopsis thalia
24	30	83.3	157	24	ABU02189	S. pneumoniae type
25	30	83.3	176	21	AAG39322	Arabidopsis thalia
26	30	83.3	838	21	AAG39321	Arabidopsis thalia
27	29	80.6	78	21	AAG25669	Arabidopsis thalia
28	29	80.6	82	21	AAG25668	Human colon cancer
29	29	80.6	129	22	AAG74100	Arabidopsis thalia
30	29	80.6	143	21	AAG25667	Arabidopsis thalia
31	29	80.6	198	21	AAG34144	Zea mays protein f
32	29	80.6	202	21	AAG54772	Arabidopsis thalia
33	29	80.6	289	21	AAG59885	Arabidopsis thalia
34	29	80.6	312	21	AAG34143	Zea mays protein f
35	29	80.6	371	21	AAG34142	Zea mays protein f
36	29	80.6	532	22	ABG19288	Novel human diagno
37	29	80.6	696	24	ABJ37047	Human breast cance
38	29	80.6	707	22	ABG19287	Novel human diagno
39	29	80.6	724	22	ABP61085	Drosophila melanog
40	29	80.6	730	24	ABP96205	Human mature perio
41	29	80.6	747	24	ABP96202	Human pericostin-L
42	29	80.6	750	24	ABP96207	Human mature TCG1
43	29	80.6	758	24	ABP96206	Human mature perio
44	29	80.6	771	24	ABP96204	Human TCG1 protein
45	29	80.6	779	14	AAR41868	Human OSF-2. Homo

ALIGNMENTS

RESULT 1

AAE14553

ID AAE14553 standard; peptide; 7 AA.

XX AAE14553;

AC AAE14553;

DT 17-MAY-2002 (first entry)

XX Human alpha-synuclein aggregation inhibitor #8.

DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;

XX Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;

XX multiple system atrophy; Hallervorden-Spatz disease; human.

OS Homo-sapiens.

XX WO200204482-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-US21379.

XX 07-JUL-2000; 2000US-217319P.

XX 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

XX Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

XX Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
 PT determining aggregation of alpha synuclein in the presence of exogenous
 PT iron or copper
 XX
 PS Claim 40; Page 37; 52pp; English.

XX The invention relates to screening of inhibitors of alpha-synuclein
 CC aggregation in the presence of exogenous iron or copper. The inhibitors
 CC are magnesium and alpha-synuclein binding peptides, which are
 CC useful for treating neurodegenerative disease that involves
 CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
 CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
 CC atrophy and Hallervorden-Spatz disease. The present sequence is a
 CC peptide that binds to the NAC (non-amyloid-beta protein component)
 CC portion of human alpha-synuclein and inhibits its aggregation.
 XX

SQ Sequence 7 AA;
 Query Match 100.0%; Score 36; DB 23; Length 7;
 Best Local Similarity 100.0%; Pred. NO. 9.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RLRGRNQ 7
 |||||
 Db 1 RLRGRNQ 7

RESULT 2

ABB63409 ABB63409 standard; Protein; 533 AA.

AC ABB63409;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 17019.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL07512.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

XX Disclosure; SEQ ID NO 17019; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 533 AA;

Query Match 100.0%; Score 36; DB 22; Length 533;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RLRGRNQ 7
 |||||
 Db 503 RLRGRNQ 509

RESULT 3

ABP35037

ID ABP35037 standard; Protein; 84 AA.

XX ABP35037;

DT 08-JUL-2002 (first entry)

XX Human ORF4010 protein, SEQ ID NO:8020.

XX Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.

OS Homo sapiens.

XX WO200190366-A2.

XX 29-NOV-2001.

XX 24-MAY-2001; 2001WO-US17076.

XX 24-MAY-2000; 2000US-206690P.

XX (CURA-) CURAGEN CORP.

XX Leach MD, Shinkets RA;

XX WPI; 2002-106200/14.

XX N-PSDB; ABL79063.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -

XX Claim 10; Page 2251; 2508pp; English.

XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide

CC range of biological activities, such as cytokine, cell proliferation,
CC cell differentiation, immune modulation, haematopoiesis regulation,
CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antiinfective activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.

XX SQ Sequence 84 AA;
Query Match 91.7%; Score 33; DB 23; Length 84;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 31 RVRGRNQ 37
:|||||

RESULT 4
ABG17799
ID ABG17799 standard; Protein; 458 AA.

XX AC ABG17799;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #17790.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX FN WO200175067-A2.
XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS81986.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID No 48158; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 458 AA;

Query Match 88.9%; Score 32; DB 22; Length 458;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
Db 253 QLRGRNQ 259
:|||||

RESULT 5
ABB49607
ID ABB49607 standard; Protein; 75 AA.

XX AC ABB49607;
XX DT 05-FEB-2002 (first entry)
XX DE Listeria monocytogenes protein #2311.
XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX OS Listeria monocytogenes.
XX FN WO200177335-A2.
XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-FR01118.
XX PR 11-APR-2000; 2000FR-0004629.

XX PA (INSP) INST PASTEUR.

XX PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
PT and prevention of Listeria and related bacterial infections, and
PT related polypeptides

PS Claim 6; SEQ ID No 2312; 192pp; French.

XX The present invention relates to the genome sequence of *Listeria*

CC monocytes EGD-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in *L.*

CC monocytes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of *L. monocytogenes* and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate *L. monocytogenes*-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by *L.*

CC monocytes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 75 AA;

Query Match 86.1%; Score 31; DB 23; Length 75;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 70 RLRGRN 75

RESULT 6

AAU61300

ID AAU61300 standard; Protein; 78 AA.

XX

AC AAU61300;

XX

DT 27-FEB-2002 (first entry)

XX

DE Propionibacterium acnes immunogenic protein #22196.

XX

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.

XX

OS Propionibacterium acnes.

XX

PN WO200181581-A2.

XX

PD 01-NOV-2001.

XX

PF 20-APR-2001; 2001WO-US12865.

XX

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX

PA (CORI-) CORIXA CORP.

XX

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX

DR WPI; 2001-616774/71.

DR N-PSDB; AAS59616.

XX

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX

PS Example 1; SEQ ID No 22495; 1069pp; English.

XX

CC Sequences AAU39105-AAU68017 represent *Propionibacterium acnes* immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

CC *P. acnes* is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of *P. acnes* in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for *P. acnes* proteins. These antibodies can be used to

CC downregulate expression and activity of *P. acnes* polypeptides and

CC therefore treat *P. acnes* infections. The antibodies may also be used as

CC diagnostic agents for determining *P. acnes* presence, for example, by

CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 78 AA;

Query Match 86.1%; Score 31; DB 22; Length 78;

Best Local Similarity 85.7%; Pred. No. 34;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7

Db 47 RLRGRDQ 53

RESULT 7

ABG20695

ID ABG20695 standard; Protein; 87 AA.

XX

AC ABG20695;

XX

DT 18-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #20686.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US08631.

XX

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR N-PSDB; AAS84882.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX

PS Claim 20; SEQ ID No 51054; 103pp; English.

XX

CC The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 87 AA;

Query Match 86.1%; Score 31; DB 22; Length 87;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLGRNQ 7
Db 32 RLGRSQ 38
|||||

RESULT 8
AAG18501
ID AAG18501 standard; Protein; 131 AA.

XX AC AAG18501;

DT 17-OCT-2000 (first entry)
XX Zea mays protein fragment SEQ ID NO: 19335.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.

OS Zea mays subsp. mays.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 131;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
 |||||
 Db 8 RLRGRN 13

RESULT 9
 AAG25683
 ID AAG25683 standard; Protein; 139 AA.
 XX
 AC AAG25683;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 29845.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 139;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 8 RLRGRN 13

RESULT 10

AAG18500

ID AAG18500 standard; Protein; 145 AA.

XX

AC AAG18500;

XX

DT 17-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 19934.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.

XX

OS Zea mays subsp. mays.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

Query Match	86.1%;	Score 31;	DB 21;	Length 145;
Best Local Similarity	100.0%;	Pred. No. 65;		
Matches	6;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	RLRGRN	6	
Db	22	RLRGRN	27	
RESULT 11				
AAG25682				
ID	AAG25682 standard; Protein; 153 AA.			
XX				
AC	AAG25682;			
XX				
DT	17-OCT-2000 (first entry)			
XX				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 29844.			
XX				
KW	Protein identification; signal transduction pathway; metabolic pathway;			
KW	hybridisation assay; genetic mapping; gene expression control; promoter;			
KW	termination sequence.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	EP1033405-A2.			
XX				
PD	06-SEP-2000.			
XX				
PF	25-FEB-2000; 2000EP-0301439.			
XX				
PR	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	23-MAR-1999; 99US-0125788.			
PR	25-MAR-1999; 99US-0126264.			
PR	29-MAR-1999; 99US-0126785.			
PR	01-APR-1999; 99US-0127462.			
PR	06-APR-1999; 99US-0128234.			
PR	08-APR-1999; 99US-0128714.			
PR	16-APR-1999; 99US-0129845.			
PR	19-APR-1999; 99US-0130077.			
PR	21-APR-1999; 99US-0130449.			
PR	23-APR-1999; 99US-0130510.			
PR	23-APR-1999; 99US-0130891.			
PR	28-APR-1999; 99US-0131449.			
PR	30-APR-1999; 99US-0132048.			
PR	30-APR-1999; 99US-0132407.			
PR	04-MAY-1999; 99US-0132484.			
PR	05-MAY-1999; 99US-0132485.			
PR	06-MAY-1999; 99US-0132486.			
PR	06-MAY-1999; 99US-0132487.			
PR	07-MAY-1999; 99US-0132863.			
PR	11-MAY-1999; 99US-0134256.			
PR	14-MAY-1999; 99US-0134218.			
PR	14-MAY-1999; 99US-0134219.			
PR	14-MAY-1999; 99US-0134221.			
PR	14-MAY-1999; 99US-0134370.			
PR	18-MAY-1999; 99US-0134768.			
PR	19-MAY-1999; 99US-0134941.			
PR	20-MAY-1999; 99US-0135124.			
PR	21-MAY-1999; 99US-0135353.			
PR	24-MAY-1999; 99US-0135629.			
PR	25-MAY-1999; 99US-0136021.			
PR	27-MAY-1999; 99US-0136392.			
PR	28-MAY-1999; 99US-0136782.			
PR	01-JUN-1999; 99US-0137222.			
PR	03-JUN-1999; 99US-0137528.			
PR	04-JUN-1999; 99US-0137502.			
PR	07-JUN-1999; 99US-0137724.			
PR	08-JUN-1999; 99US-0138094.			
PR	10-JUN-1999; 99US-0138540.			
PR	10-JUN-1999; 99US-0138847.			

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 153;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 RLRGRN 6
|||
Db 22 RLRGRN 27

RESULT 12
AAG16627
ID AAG16627 standard; Protein; 181 AA.
XX
AC AAG16627;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17345.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 181;

Best Local Similarity 100.0%; Pred. No. 81;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6

Db 8 RLRGRN 13

RESULT 13

AAG16626

ID AAG16626 standard; Protein; 195 AA.

XX

AC AAG16626;

XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 17344.
DE Arabidopsis thaliana
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 195;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLRGRN 6
|||
Db 22 RLRGRN 27

RESULT 14
ABB50063
ID ABB50063 standard; Protein; 280 AA.

XX ABB50063;

XX 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #2767.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.
XX WO200177335-A2.
XX 18-OCT-2001.
XX 11-APR-2001; 2001WO-FR01118.
XX 11-APR-2000; 2000FR-0004629.
XX (INSP) INST PASTEUR.
XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Krefit J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX WPI; 2002-010914/01.
XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
XX and prevention of Listeria and related bacterial infections, and
XX related polypeptides -
XX Claim 6; SEQ ID No 2768; 192pp; French.
XX The present invention relates to the genome sequence of Listeria
XX monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
XX it are useful for selecting probes and primers for detecting genes in L.
XX monocytogenes and related organisms, and for studying genetic
XX polymorphisms and other genomes. The present sequence is a protein
XX encoded by the genome sequence of the present invention. Proteins
XX expressed from the genome sequence are useful for raising specific
XX antibodies, identification of L. monocytogenes and related organisms, and
XX for biosynthesis and biodegradation, especially biosynthesis of Vitamin
XX B12. The genome sequence and proteins encoded by it are also useful for
XX selecting compounds that regulate gene expression and cell replication
XX and modulate L. monocytogenes-related diseases. In addition, the genome
XX sequence and proteins encoded by it are useful in pharmaceutical and
XX vaccines compositions for the treatment or prevention of infections by L.
XX monocytogenes and related organisms.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 280 AA;
Query Match 86.1%; Score 31; DB 23; Length 280;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLRGRNQ 7
|||
Db 192 RLHGRNQ 198

RESULT 15
AAU46599
ID AAU46599 standard; Protein; 337 AA.
XX AC AAU46599;
XX DT 27-FEB-2002 (first entry)
XX DE Propionibacterium acnes immunogenic protein #7495.
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;

KW dermatological; osteopathic; neuroprotectant.
XX Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US12865.
XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX (CORI-) CORIXA CORP.
PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59534.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX Example 1; SEQ ID No 7794; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 337 AA;

Query Match 86.1%; Score 31; DB 22; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRGRNQ 7
|||
Db 136 LRGRNQ 141

Search completed: February 11, 2004, 17:02:52
Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	196	5 Q9U3C2	Q9u3c2 caenorhabdi
2	42	79.2	272	13 Q8JH73	Q8jh73 xenopus lae
3	42	79.2	309	10 Q49746	Q49746 arabidopsis
4	42	79.2	310	10 Q9M2Y9	Q9m2y9 arabidopsis
5	42	79.2	487	16 Q9CKV7	Q9ckv7 pasteurella
6	42	79.2	2197	5 Q9W016	Q9w016 drosophila
7	40	75.5	732	13 Q9GU61	Q9gu61 myxine glut
8	40	75.5	749	11 Q8BQQ7	Q8bqq7 mus musculu
9	40	75.5	1826	4 Q8IZX4	Q8izz4 homo sapien
10	40	75.5	1865	11 Q60544	Q60544 mesocricetu
11	39	73.6	179	11 Q9CUD2	Q9cud2 mus musculu
12	39	73.6	242	10 Q9ARQ2	Q9arq2 oryza sativ
13	39	73.6	375	6 Q97505	Q97505 sus scrofa
14	39	73.6	563	2 Q9RS37	Q9rs37 pseudomonas
15	39	73.6	750	10 Q81880	Q81880 arabidopsis
16	39	73.6	826	10 Q940X9	Q940x9 arabidopsis

17	38.5	72.6	750	3	O42657	O42657 schizosacch
18	38	71.7	193	10	Q9LWM8	Q9lwm8 oryza sativ
19	38	71.7	284	16	O06252	O06252 mycobacteri
20	38	71.7	293	3	Q12393	Q12393 saccharomyc
21	38	71.7	312	16	Q8VJ08	Q8vj08 mycobacteri
22	38	71.7	316	16	Q9KS69	Q9ks69 vibrio chol
23	38	71.7	370	10	Q9LGV4	Q9lgv4 oryza sativ
24	38	71.7	381	8	Q8HHQ4	Q8hhq4 sigmodon hi
25	38	71.7	412	10	Q8RYL7	Q8ryl7 oryza sativ
26	38	71.7	545	2	Q9F3V6	Q9f3v6 pseudonocar
27	38	71.7	674	5	Q93550	Q93550 caenorhabdi
28	38	71.7	708	16	Q8CZY0	Q8czy0 yersinia pe
29	38	71.7	753	16	Q8ZHU0	Q8zhu0 yersinia pe
30	38	71.7	766	16	O06817	O06817 mycobacteri
31	37.5	70.8	79	5	Q94188	Q94188 caenorhabdi
32	37.5	70.8	452	4	Q96AB7	Q96ab7 homo sapien
33	37.5	70.8	484	4	Q9BTV6	Q9btv6 homo sapien
34	37	69.8	100	12	Q9YRJ4	Q9yrj4 human cytom
35	37	69.8	101	12	Q9YRI7	Q9yri7 human cytom
36	37	69.8	105	12	Q9YRJ6	Q9yrj6 human cytom
37	37	69.8	111	12	Q9YRJ8	Q9yrj8 human cytom
38	37	69.8	112	12	Q9YRJ7	Q9yrj7 human cytom
39	37	69.8	119	9	Q8LTE5	Q8lte5 cyanophage
40	37	69.8	138	12	Q9YRI5	Q9yri5 human cytom
41	37	69.8	139	12	Q9YRJ2	Q9yrj2 human cytom
42	37	69.8	142	10	Q94J29	Q94j29 oryza sativ
43	37	69.8	221	3	Q9P8K9	Q9p8k9 emeritella
44	37	69.8	299	10	Q8RWG4	Q8rwg4 arabidopsis
45	37	69.8	313	10	Q9M0M0	Q9m0m0 arabidopsis

ALIGNMENTS

RESULT 1
Q9U3C2
ID Q9U3C2; PRELIMINARY; PRT; 196 AA.
AC Q9U3C2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE K03H1.9 protein.
GN K03H1.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith M.J., Ainscough R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z29560; CAA82661.2; -.
DR WormPep; K03H1.9; CE23851.
SQ SEQUENCE 196 AA; 24273 MW; 958E7B149C150D4C CRC64;

Query Match 79.2%; Score 42; DB 5; Length 196;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFFHHR 7
| | | | |
Db 78 PFFHHR 83

RESULT 2
Q8JH73

```
ID Q8JH73 PRELIMINARY; PRT; 272 AA.
AC Q8JH73;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Integral membrane nucleoporin gp210 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22100007; PubMed=12093788;
RA Drummond S.P., Wilson K.L.;
RT "Interference with the cytoplasmic tail of gp210 disrupts 'close
RT apposition' of nuclear membranes and blocks nuclear pore dilation.";
RL J. Cell Biol. 158:53-62(2002).
DR EMBL; AF533550; AAM94631.1; -.
KW Porin.
FT NON_TER 1 1
SQ SEQUENCE 272 AA; 29952 MW; EE9023786F0F9988 CRC64;

Query Match 79.2%; Score 42; DB 13; Length 272;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PPFHHR 7
Db 225 PPFHHR 230

RESULT 3
O49746 PRELIMINARY; PRT; 309 AA.
ID O49746;
AC O49746;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE R2R3-MYB transcription factor.
GN MYB84.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX Romero I., Fustes A., Benito M.J., Malpica J., Leyva A., Paz-Ares J.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL; Y14209; CAA74605.1; -.
DR EMBL; Z95807; CAB09239.1; -.
DR HSSP; P06876; 1MBK.
DR TRANSFAC; T02597; -.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 309 AA; 35545 MW; 8E673F408803007C CRC64;

Query Match 79.2%; Score 42; DB 10; Length 309;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 4
O9M2Y9 PRELIMINARY; PRT; 310 AA.
ID O9M2Y9;
AC O9M2Y9;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE AtMYB84.
GN T16K5.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX EU Arabidopsis sequencing project;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 2 MYB-LIKE DOMAINS.
DR EMBL; ALI32965; CAB66907.1; -.
DR HSSP; P06876; 1MBK.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; myb_DNA-binding; 2.
DR SMART; SM00717; SANT; 2.
DR PROSITE; PS00037; MYB_1; 1.
DR PROSITE; PS00334; MYB_2; 1.
DR PROSITE; PS50090; MYB_3; 2.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 310 AA; 35578 MW; 647FB85D2226CB6C3 CRC64;

Query Match 79.2%; Score 42; DB 10; Length 310;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 5
O9CKV7 PRELIMINARY; PRT; 487 AA.
ID O9CKV7;
AC O9CKV7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE TrkH.
GN TRKH OR PM1498.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Pm70;
RC MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006186; AAK03582.1; -.
DR InterPro; IPR003445; Cat_transpt.
DR InterPro; IPR004772; K_transptTrk.
DR Pfam; PF02386; TrkH; 1.
DR TIGRFAMs; TIGR00933; 2a38; 1.
KW Complete proteome.
SQ SEQUENCE 487 AA; 54000 MW; 15049745C5872BDE CRC64;
```


Query Match 75.5%; Score 40; DB 13; Length 732;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 53 WPFHH 57

RESULT 8

Q8BQQ7 PRELIMINARY; PRT; 749 AA.
AC Q8BQQ7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TAFII250 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK046668; BAC32828.1; --
FT NON TER 1
SQ SEQUENCE 749 AA; 85475 MW; 2FAF7354AEC42710 CRC64;

Query Match 75.5%; Score 40; DB 11; Length 749;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 405 WPFHH 409

RESULT 9

Q81ZX4 PRELIMINARY; PRT; 1826 AA.
AC Q81ZX4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TBP-associated factor RNA polymerase 1-like.
GN TAF1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22206412; PubMed=12217962;
RA Wang P.J., Page D.C.;
RT "Functional substitution for TAF(II)250 by a retroposed homolog that
is expressed in human spermatogenesis."
RL Hum. Mol. Genet. 11:2341-2346 (2002).
DR EMBL; AF390562; AAN40840.1; --
SQ SEQUENCE 1826 AA; 207300 MW; 35D780E749AC9B17 CRC64;

Query Match 75.5%; Score 40; DB 4; Length 1826;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 405 WPFHH 409

Db 1545 WPFHH 1549

RESULT 10

Q60544 PRELIMINARY; PRT; 1865 AA.
AC Q60544;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CCG1.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94215915; PubMed=8163200;
RA Hayashida T., Sekiguchi T., Noguchi E., Sunamoto H., Ohba T.,
RA Nishimoto T.;
RT "The CCG1/TAFII250 gene is mutated in thermosensitive G1 mutants of
the BHK21 cell line derived from goldenhamster."
RL Gene 141:267-270 (1994).
DR EMBL; D26114; BAA05110.1; --
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
SQ SEQUENCE 1865 AA; 211866 MW; A81614946C0C0F24 CRC64;

Query Match 75.5%; Score 40; DB 11; Length 1865;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 1521 WPFHH 1525

RESULT 11

Q9CUD2 PRELIMINARY; PRT; 179 AA.
AC Q9CUD2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 4930520K10Rik protein (Fragment).
GN 4930520K10RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

```
RA  Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA  Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA  Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA  Hayashizaki Y.,
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
DR  EMBL; AK016752; BAB30410.1; -.
DR  MGD; MGI:1914837; 4930520K10Rik.
FT  NON_TER
SQ  SEQUENCE 179 AA; 18465 MW; 583C70B92085ECE9 CRC64;

Query Match 73.6%; Score 39; DB 11; Length 179;
Best Local Similarity 71.4%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 7
Db 6 WKFHHR 12

RESULT 12
Q9ARQ2 PRELIMINARY; PRT; 242 AA.
AC Q9ARQ2; (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE OSJNBa0010K01.13 protein.
GN OSJNBa0010K01.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:OSJNBa0010K01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003210; BAB40099.1; -.
DR Gramene; Q9ARQ2; -.
DR InterPro; IPR007087; Znf_C2H2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
SQ SEQUENCE 242 AA; 27507 MW; EAE4BA06AA06D785 CRC64;

Query Match 73.6%; Score 39; DB 10; Length 242;
Best Local Similarity 83.3%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 51 WPGHHH 56

RESULT 13
O97505 PRELIMINARY; PRT; 375 AA.
AC O97505;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuropeptide Y receptor type 4 (Neuropeptide Y-family receptor
DE Y4).
GN NPY Y4.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LWD; TISSUE=Kidney;
```

```
RA Ito Y., Minezawa M.;
RT "Sus scrofa NPY Y4 gene for neuropeptide Y receptor type 4, complete
RT cds [genomic].";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BREED NORWEGIAN LANDRACE;
RX MEDLINE=20187975; PubMed=10720571;
RA Wraith A., Tornsten A., Chardon P., Harbitz I., Chowdhary B.P.,
RA Andersson L., Lundin L.-G., Larhammar D.;
RT "Evolution of the neuropeptide Y receptor family: gene and chromosome
RT duplications deduced from the cloning and mapping of the five receptor
RT subtype genes in the pig.";
RL Genome Res. 10:302-310(2000).
DR EMBL; AB021678; BAA36218.1; -.
DR EMBL; AF227955; AAF62507.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Neuropeptide; Receptor.
SQ SEQUENCE 375 AA; 42346 MW; 5992F075904DB1A0 CRC64;

Query Match 73.6%; Score 39; DB 6; Length 375;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHH 7
Db 205 WPLEHHR 211

RESULT 14
Q8RS37 PRELIMINARY; PRT; 563 AA.
AC Q8RS37;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE WD-repeat protein.
GN TOXC.
OS Pseudomonas glumae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=337;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF327195;
RA Suzuki F.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF327195;
RA Suzuki F., Sawada H., Matsuda I.;
RT "Molecular Characterization of Toxoflavin Biosynthesis-related Gene in
RT Pseudomonas (Burkholderia) glumae.";
RL Ann. Phytopathol. Soc. Jpn. 64:276-281(1998).
DR EMBL; AB040403; BAB88914.1; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 563 AA; 61957 MW; 6636126F0590931D CRC64;

Query Match 73.6%; Score 39; DB 2; Length 563;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHH 7
```

Db 501 WPLHRH 507

RESULT 15
O81880
ID O81880 PRELIMINARY; PRT; 750 AA.
AC O81880;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical 83.9 kDa protein.
GN T16L1.120 OR AT4G33630.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Deutschenbaur S., Piravandi E., Hohelsel J., Jesse T.,
RA Heijnen L., Vos P., Mewes H.W., Mayer K.F.X., Schueller C., Bevan M.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Obermaier B., Deutschenbaur S., Piravandi E., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031394; CAA20576.1; -
DR EMBL; AL161583; CAB80080.1; -
KW Hypothetical protein.
SQ SEQUENCE 750 AA; 83907 MW; 3F298ECB1FD505A6 CRC64;

Query Match 73.6%; Score 39; DB 10; Length 750;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHR 7
|:||||
Db 21 PYHHR 26

Search completed: February 11, 2004, 17:09:41
Job time : 33.5833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	79.2	487	1	TRKH_HAEIN
2	40	75.5	1872	1	T2D1_HUMAN
3	39	73.6	316	1	COAA_BACHD
4	38	71.7	744	1	BISC_RHOSH
5	38	71.7	925	1	PIP1_YEAST
6	37	69.8	221	1	HEX1_EMENI
7	37	69.8	236	1	MRAY_ARATH
8	37	69.8	327	1	COL7_ARATH
9	37	69.8	336	1	MRAY_STRP3
10	37	69.8	336	1	MRAY_STRP8
11	37	69.8	336	1	MRAY_STRPY
12	37	69.8	359	1	MRAY_AQUAE
13	37	69.8	359	1	MRAY_MYCLE
14	37	69.8	359	1	MRAY_MYCTU
15	37	69.8	375	1	NY4R_HUMAN
16	37	69.8	395	1	DLTB_BACSU
17	37	69.8	611	1	THIC_RHIME
18	37	69.8	619	1	THIC_RHILO
19	37	69.8	1849	1	T172_HUMAN
20	36	67.9	404	1	ZFS1_SCHPO
21	36	67.9	526	1	CNA2_YEAST
22	35	66.0	254	1	S5A2_HUMAN
23	35	66.0	254	1	S5A2_MACFA
24	35	66.0	254	1	S5A2_PIG
25	35	66.0	319	1	TCB1_RABIT
26	35	66.0	356	1	POU1_XENLA
27	35	66.0	386	1	HXAD_MOUSE
28	35	66.0	388	1	HXAD_HUMAN
29	35	66.0	447	1	HEX_ADE04
30	35	66.0	457	1	THIC_ANASP
31	35	66.0	459	1	THIC_SYNY3
32	35	66.0	465	1	HEX_ADE06
33	35	66.0	467	1	HEX_ADE01

34	35	66.0	468	1	HEX_ADE31	P36855 human adeno
35	35	66.0	501	1	FXC2_HUMAN	Q99958 homo sapien
36	35	66.0	512	1	HEX_ADE09	P36853 human adenc
37	35	66.0	517	1	HEX_ADE08	P36852 human adeno
38	35	66.0	905	1	HEX_ADECC	Q65955 canine aden
39	35	66.0	905	1	HEX_ADECR	O39619 canine aden
40	35	66.0	909	1	HEX_ADEM1	P48308 mouse adeno
41	35	66.0	911	1	HEX_ADEB3	P03278 bovine aden
42	35	66.0	919	1	HEX_ADE12	P19900 human adeno
43	35	66.0	923	1	HEX_ADE40	P11819 human adeno
44	35	66.0	925	1	HEX_ADE41	P11820 human adeno
45	35	66.0	937	1	HEX_ADE07	P36851 human adeno

ALIGNMENTS

RESULT 1

TRKH_HAEIN ID	TRKH_HAEIN	STANDARD;	PRT;	487 AA.
AC	P44843;	01-NOV-1995 (Rel. 32, Created)		
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Trk system potassium uptake protein trkh.			
GN	TRKH OR HI0723.			
OS	Haemophilus influenzae.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;			
OC	Pasteurellaceae; Haemophilus.			
OX	NCBI_TaxID=727;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=rd / KW20 / ATCC 51907;			
RX	MEDLINE=95350630; PubMed=7542800;			
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,			
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,			
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,			
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,			
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,			
RA	Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,			
RA	Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,			
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,			
RA	Venter J.C.;			
RT	"Whole-genome random sequencing and assembly of Haemophilus influenzae			
RL	Rd.";			
RL	Science 269:496-512(1995).			
RN	[2]			
RP	REVISIONS.			
RA	White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D.,			
RA	Peterson J., Hickey E., Dodson R., Gwinn M.;			
RL	Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: LOW-AFFINITY POTASSIUM TRANSPORT SYSTEM. INTERACTS WITH			
CC	TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA AND REQUIRES TRKE FOR			
CC	TRANSPORT ACTIVITY (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(Potential).			

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U32755; AAC22381.1; -.
TIGR; HI0723; -.
DR InterPro; IPR003445; Cat_transpt.
DR InterPro; IPR004772; K_transptTrk.
DR Pfam; PF02386; Trkh; 1.
DR TIGRFAMs; TIGR00933; 2a38; 1.
DR Transmembrane; Inner membrane; Transport; Potassium transport;
KW Complete proteome.

```
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 280 300 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 401 421 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
FT TRANSMEM 461 481 POTENTIAL.
SQ SEQUENCE 487 AA; 53468 MW; 91A8F38C37F4540A CRC64;

Query Match 79.2%; Score 42; DB 1; Length 487;
Best Local Similarity 71.4%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
Db 55 WPCHHK 61

RESULT 2
T2D1_HUMAN
ID T2D1_HUMAN STANDARD; PRT; 1872 AA.
AC P21675;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Transcription initiation factor TFIID 250 kDa subunit (TAFII-250)
DE (TAFII250) (TBP-associated factor 250 kDa) (P250) (Cell cycle gene 1
DE protein).
GN TAF1 OR TAF2A OR CCG1 OR BA2R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Laryngeal carcinoma;
RX MEDLINE=91246200; PubMed=2038334;
RA Sekiguchi T., Nohiro Y., Nakamura Y., Hisamoto N., Nishimoto T.;
RT "The human CCG1 gene, essential for progression of the G1 phase,
RT encodes a 210-kilodalton nuclear DNA-binding protein.";
RL Mol. Cell. Biol. 11:3317-3325(1991).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RX MEDLINE=89005056; PubMed=3169001;
RA Sekiguchi T., Miyata T., Nishimoto T.;
RT "Molecular cloning of the cDNA of human X chromosomal gene (CCG1)
RT which complements the temperature-sensitive G1 mutants, tsBN462 and
RT ts13, of the BHK cell line.";
RL EMBO J. 7:1683-1687(1988).
RN [3]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=93196704; PubMed=7680771;
RA Ruppert S., Wang E.H., Tjian R.;
RT "Cloning and expression of human TAFII250: a TBP-associated factor
RT implicated in cell-cycle regulation.";
RL Nature 362:175-179(1993).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=93196705; PubMed=8450888;
RA Hisatake K., Hasegawa S., Takada R., Nakatani Y., Horikoshi M.,
RA Roeder R.G.;
RT "The p250 subunit of native TATA box-binding factor TFIID is the
RT cell-cycle regulatory protein CCG1.";
RL Nature 362:179-181(1993).
CC -!- FUNCTION: MAY PLAY AN ESSENTIAL ROLE IN TFIID ASSEMBLY BY
CC INTERACTING WITH BOTH TBP AND OTHER TAF, AS WELL AS SERVING TO
CC LINK THE CONTROL OF TRANSCRIPTION TO THE CELL CYCLE. ESSENTIAL
CC FOR PROGRESSION OF THE G1 PHASE OF THE CELL CYCLE. POSSESSES
```

DNA-BINDING ACTIVITY.
-!- SUBUNIT: TFIID IS COMPOSED OF TBP AND A VARIETY OF TBP-ASSOCIATED FACTORS.
-!- SUBCELLULAR LOCATION: Nuclear.
-!- PTM: PHOSPHORYLATED BY CASEIN KINASE II IN VITRO.
-!- SIMILARITY: Contains 2 bromodomains.
-!- SIMILARITY: Contains 1 HMG box domain.
-!- SIMILARITY: TO DROSOPHILA TAFII-230. SOME TO S.POMBE TAFII-111 AND TO S.CEREVISIAE TAFI45.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D90359; BAA14374.1; --
EMBL; X07024; CAA30073.1; ALT_SEQ.
PIR; A40262; A40262.
PDB; 1EQF; 07-JUN-00.
TRANSFAC; T02206; --
Gene; HGNC:11535; TAF1.
MIM; 313650; --
GO; GO:0005669; C:transcription factor TFIID complex; TAS.
GO; GO:0016251; F:general RNA polymerase II transcription fac. . .; TAS.
GO; GO:0004672; F:protein kinase activity; TAS.
GO; GO:0000114; P:G1-specific transcription in mitotic cell c. . .; TAS.
GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
GO; GO:0006461; P:protein complex assembly; TAS.
InterPro; IPR001487; Bromodomain.
Pfam; PF00439; bromodomain; 2.
PRINTS; PR00503; BROMODOMAIN.
SMART; SM00297; BROMO; 2.
PROSITE; PS00633; BROMODOMAIN_1; 2.
PROSITE; PS0014; BROMODOMAIN_2; 2.
Bromodomain; Nuclear protein; DNA-binding; Cell cycle; Repeat;
Transcription regulation; Phosphorylation; 3D-structure.
DOMAIN 157 165 PRO-RICH.
DNA BIND 1195 1273 HMG BOX (POTENTIAL).
DOMAIN 1351 1358 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
DOMAIN 1397 1467 BROMODOMAIN 1.
DOMAIN 1520 1590 BROMODOMAIN 2.
DOMAIN 1627 1872 ASP/GLU-RICH (ACIDIC TAIL).
SEQUENCE 1872 AA; 212676 MW; 93BE3D181A72ABEB CRC64;
Query Match 75.5%; Score 40; DB 1; Length 1872;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WPFHH 5
Db 1526 WPFHH 1530

RESULT 3
COAA_BACHD
ID COAA_BACHD STANDARD; PRT; 316 AA.
AC Q9K8X7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pantothenate kinase (EC 2.7.1.33) (Pantothenic acid kinase).
GN COAA OR BH2875.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hiram C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*
RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + pantothenate = ADP + D-4'-
CC phosphopantothenate.
CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC PANTOTHENATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP001516; BAB06594.1; -.
CC PIR; C84009; C84009.
CC HAMAP; MF 00215; -; 1.
CC InterPro; IPR004566; Pank_bact.
CC InterPro; IPR006083; PRK_URK.
CC Pfam; PF00485; PRK; 1.
CC TIGRFAMs; TIGR00554; pank_bact; 1.
CC Transferase; Kinase; ATP-Binding; Coenzyme A biosynthesis;
KW Complete proteome.
FT NP_BIND 95 102 ATP (POTENTIAL).
SQ SEQUENCE 316 AA; 36416 MW; 3196BA013E3B0BB3 CRC64;

Query Match 73.6%; Score 39; DB 1; Length 316;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
DB 115 WPEHHH 120

RESULT 4
BISC_RHOSH STANDARD; PRT; 744 AA.
AC P54934;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Biotin sulfoxide reductase (EC 1.1.1.1) (BDS reductase) (BSO
DE reductase).
OS Rhodobacter sphaeroides (Rhodospseudomonas sphaeroides).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. sp. denitrificans IL106;
RX MEDLINE=95251380; PubMed=7733660;
RA Pollock V.V., Barber M.J.;
RT "Molecular cloning and expression of biotin sulfoxide reductase from
RT Rhodobacter sphaeroides forma sp. denitrificans.";
RL Arch. Biochem. Biophys. 318:322-332(1995).
CC -1- FUNCTION: THIS ENZYME MAY SERVE AS A SCAVENGER, ALLOWING THE CELL
CC TO UTILIZE BIOTIN SULFOXIDE AS A BIOTIN SOURCE. IT REDUCES A
CC SPONTANEOUS OXIDATION PRODUCT OF BIOTIN, D-BIOTIN D-SULFOXIDE (BSO
CC OR BDS), BACK TO BIOTIN (BY SIMILARITY).
CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U08189; AAA74739.1; -.
CC PIR; S65669; S65669.
CC HSSP; Q57366; 1EU1.
CC InterPro; IPR006658; Bisc.
CC InterPro; IPR006657; Mol_dinuc_bind.
CC InterPro; IPR006656; Molybdopterin.
CC InterPro; IPR006655; Prok_Mboxred.
CC Pfam; PF00384; molybdopterin; 1.
CC Pfam; PF01568; Molybdop_binding; 1.
CC TIGRFAMs; TIGR00509; bisc_fam; 1.
CC PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; FALSE_NEG.
CC PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
CC PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
KW Oxidoreductase; Molybdenum.
SQ SEQUENCE 744 AA; 80266 MW; 731A933E95358A19 CRC64;

Query Match 71.7%; Score 38; DB 1; Length 744;
Best Local Similarity 83.3%; Pred. No. 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHHR 7
DB 401 PFHHHQ 406

RESULT 5
PIPL_YEAST STANDARD; PRT; 925 AA.
AC P40020;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymerase-interacting protein 1 (Factor interacting with REP).
DE PIPI OR FIR1 OR YER032W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RA del Olmo M., Gross S., Moore C.L.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX PubMed=9169868;
RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V.";
RL Nature 387:78-81(1997).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=97339480; PubMed=9196079;
RA Russnak R., Pereira S., Platt T.;
RT "RNA binding analysis of yeast REF2 and its two-hybrid interaction
RT with a new gene product, FIR1.";
RL Gene Expr. 6:241-258(1996).
CC -1- FUNCTION: Interacts with poly(A) polymerase and with REF2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U17262; AAB46625.1; -;
DR EMBL; U18778; AAB64565.1; -;
DR PIR; S50490; S50490.
DR SGD; S0000834; FIR1.
DR GO; GO:0006378; P:mrna polyadenylation; IGI.
FT CONFLICT 663 663 R -> P (IN REF. 2).
SQ SEQUENCE 925 AA; 104701 MW; 707D9839EE31322B CRC64;

Query Match 71.7%; Score 38; DB 1; Length 925;
Best Local Similarity 71.4%; Pred. No. 75;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WPFHHH 7
Db 179 WKFHNR 185

RESULT 6
HEX1_EMENI STANDARD; PRT; 221 AA.
AC Q9P8K9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Woronin body major protein.
GN HEX1.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=162425;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20245872; PubMed=10783241;
RA Jedd G., Chua N.-H.;
RT "A new self-assembled peroxisomal vesicle required for efficient
resealing of the plasma membrane."
RL Nat. Cell Biol. 2:226-231(2000).
CC -!- FUNCTION: Major protein in the hexagonal crystals of Woronin
bodies, a peroxisomal vesicle that seal the septal pore in
response to cellular damage.
CC -!- SUBUNIT: Forms oligomers. Self-assemble into hexagonal rods (By
similarity).
CC -!- SUBCELLULAR LOCATION: Peroxisomal. Woronin bodies.
CC -!- SIMILARITY: Belongs to the eIF-5A family.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE239659; AAF67173.1; -;
DR InterPro; IPR001884; EIF5A_hypusine.
DR Pfam; PF01287; eIF-5a; 1.
KW Peroxisome.
FT SITE 219 221 MICROBODY TARGETING SIGNAL (POTENTIAL).
SQ SEQUENCE 221 AA; 25091 MW; 567634BFA7A62C93 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 221;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PFFHH 6
Db 27 PFFHH 31

RESULT 7
MRAY_ARATH STANDARD; PRT; 236 AA.
AC O49730;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase homolog.
GN AT4G18270 OR T9A21.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
Harris B., Ansoirge W., Brandt P., Grivell L.A., Rieger M., Mueller M.,
Weichselgartner M., de Simone V., Obermaier B., Mache R., Schmidheini T.,
Kreiss M., Delsen M., Puigdomenech P., Watson M., Schmidt M.,
Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
Langham S.-A., McCullagh B., Bilham L., Robben J.,
Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
Weitzenecker T., Bothe G., Ransperger U., Hilbert H., Braun M.,
Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
Clark I., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
Borkova D., Bloeker H., Scharfe M., Grimm M., Loehnert T.-H.,
Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,
Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
Schnabl S., Hiller R., Schmidt W., Lechharny A., Aubourg S.,
Chedford F., Cooke R., Berger C., Monfort A., Casacuberta E.,
Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
Kramer J., Fulton L., Mardis E., Dente M., Pepin K., Hillier L.,
Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
thaliana."
RL Nature 402:769-777(1999).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC -!- CAUTION: IT IS PROBABLE THAT THIS IS ONLY A FRAGMENT OF THE
CC COMPLETE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AL021713; CAA16799.1; -.
DR PIR; T04929; T04929.
DR InterPro; IPR000715; Glyco_trans 4.
DR InterPro; IPR003524; PNACpPept_trans.
DR Pfam; PF00953; Glycos_transf_4; 1.
DR PROSITE; PS01347; MRAY_1; 1.
DR PROSITE; PS01348; MRAY_2; 1.
KW Hypothetical protein; Transferase; Transmembrane.
FT TRANSMEM 23 43 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 213 233 POTENTIAL.
SQ SEQUENCE 236 AA; 24830 MW; D93F9FC4A86E3D28 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHH 6
Db 198 PFHHH 202

RESULT 8
COL7 ARATH STANDARD; PRT; 327 AA.
AC Q9SYM2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Zinc finger protein constans-like 7.
GN AT1G78600 OR T30F21.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziali A.,
RA Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.
CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AC007260; AAD30576.1; -.
DR PIR; F96814; F96814.
DR InterPro; IPR000315; Znf_Box.
DR InterPro; IPR002926; Znf_constans.
DR Pfam; PF00643; zf-B_box; 2.
DR ProDom; PD007661; Znf_constans; 2.
DR SMART; SM00336; BBOX; 2.
DR PROSITE; PS01119; ZF_BOX; 2.
KW Zinc-finger; Nuclear protein; Repeat; Multigene family.
FT ZN_FING 5 47 B_BOX-TYPE 1.
FT ZN_FING 57 99 B_BOX-TYPE 2.
FT DOMAIN 149 155 POLY-GLN.
SQ SEQUENCE 327 AA; 35932 MW; C9683E444D0504D6 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 327;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 WP--FHGH 6
Db 291 WPKHFFHH 298

RESULT 9
MRAY STRP3 STANDARD; PRT; 336 AA.
ID MRAY STRP3 STANDARD; PRT; 336 AA.
AC O8K6C7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
MURNAc-pentapeptide phosphotransferase).
GN MRAY OR SPYM3_1400 OR SPS0462.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.P., Perkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schliefert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
phage-encoded toxins, the high-virulence phenotype, and clone
emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1 / Serotype M3;
RA Nakagawa I., Kurokawa K., Nakata M., Tomiyasu Y., Yamashita A.,
RA Yamazaki K., Okahashi N., Kawabata S., Yasunaga T., Hattori M.,
RA Hayashi H., Hamada S.;
RT "The genome of invasive Streptococcus pyogenes; a comparative analysis
of S. pyogenes SSI-1, SF370 and MGAS8232.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE014163; AAM80007.1; -.
CC EMBL; AP005142; BAC63557.1; -.
CC HAMAP; MF 00038; -.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACpPep_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; FALSE_NEG.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 47 69 POTENTIAL.
FT TRANSMEM 79 98 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT TRANSMEM 143 165 POTENTIAL.
FT TRANSMEM 178 197 POTENTIAL.
FT TRANSMEM 202 219 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 252 274 POTENTIAL.
FT TRANSMEM 316 335 POTENTIAL.
SQ SEQUENCE 336 AA; 36886 MW; 7C8D20407E7B7B92 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PFHHH 6
Db 292 PFHHH 296

RESULT 10
MRAY_STRP8 STANDARD; PRT; 336 AA.
AC Q8NZY2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR SPYM18 1673.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.

```

```

CC -----
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010079; AAL98215.1; -.
CC HAMAP; MF 00038; -.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACpPep_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; FALSE_NEG.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 47 69 POTENTIAL.
FT TRANSMEM 79 98 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT TRANSMEM 143 165 POTENTIAL.
FT TRANSMEM 178 197 POTENTIAL.
FT TRANSMEM 202 219 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 252 274 POTENTIAL.
FT TRANSMEM 316 335 POTENTIAL.
SQ SEQUENCE 336 AA; 37075 MW; 6BFD8E55D71F67E2 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PFHHH 6
Db 292 PFHHH 296

RESULT 11
MRAY_STRPY STANDARD; PRT; 336 AA.
AC Q99YK2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR SPY1662.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.

```

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006597; AAK34425.1; -
CC HAMAP; MF_00038; -; 1.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACPept_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; FALSE_NEG.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
KW TRANSMEM 4 26 POTENTIAL.
FT TRANSMEM 47 69 POTENTIAL.
FT TRANSMEM 79 98 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT TRANSMEM 143 165 POTENTIAL.
FT TRANSMEM 178 197 POTENTIAL.
FT TRANSMEM 202 219 POTENTIAL.
FT TRANSMEM 226 248 POTENTIAL.
FT TRANSMEM 252 274 POTENTIAL.
FT TRANSMEM 316 335 POTENTIAL.
SQ SEQUENCE 336 AA; 36924 MW; 093C5B79CC0510CD CRC64;
Query Match 69.8%; Score 37; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
Db 292 PFHHH 296
RESULT 12
MRAY_AQUAE STANDARD; PRT; 359 AA.
AC O66465;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR AQ_053.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000671; AAC06418.1; -
CC PIR; F70304; F70304.
CC HAMAP; MF_00038; -; 1.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACPept_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; 1.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
KW TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 198 218 POTENTIAL.
FT TRANSMEM 234 254 POTENTIAL.
FT TRANSMEM 261 281 POTENTIAL.
FT TRANSMEM 287 307 POTENTIAL.
FT TRANSMEM 336 356 POTENTIAL.
SQ SEQUENCE 359 AA; 40341 MW; B2D6292EA6EA16D4 CRC64;
Query Match 69.8%; Score 37; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
Db 322 PFHHH 326
RESULT 13
MRAY_MYCLE STANDARD; PRT; 359 AA.
AC O69555;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR MURX OR ML0911 OR MLCB268.05C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus";
RL Nature 409:1007-1011(2001).
CC -!- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC

CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL022602; CAA18671.1; -.
CC EMBL; AL583920; CAC31292.1; -.
CC PIR; A87023; A87023.
CC Leproma; ML0911; -.
CC HAMAP; MF_00038; -; 1.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACpPpt_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; 1.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
CC Complete proteome.
CC TRANSMEM 3 23 POTENTIAL.
CC TRANSMEM 55 75 POTENTIAL.
CC TRANSMEM 80 100 POTENTIAL.
CC TRANSMEM 117 137 POTENTIAL.
CC TRANSMEM 156 176 POTENTIAL.
CC TRANSMEM 187 207 POTENTIAL.
CC TRANSMEM 231 251 POTENTIAL.
CC TRANSMEM 255 275 POTENTIAL.
CC TRANSMEM 280 300 POTENTIAL.
CC TRANSMEM 334 354 POTENTIAL.
CC SEQUENCE 359 AA; 37773 MW; 15554A306A3B4429 CRC64;
Query Match 69.8%; Score 37; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
DB 315 PFHHH 319
RESULT 14
MRAY_MYCTU STANDARD; PRT; 359 AA.
ID MRAY_MYCTU
AC O06221;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNAc-pentapeptide phosphotransferase).
GN MRAY OR MURX OR RV2156C OR MT2215 OR MTCY270.12.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whithead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z95388; CAB08671.1; -.
CC EMBL; AE007068; AAK46499.1; -.
CC PIR; H70579; H70579.
CC TIGR; MT2215; -.
CC TubercuList; RV2156C; -.
CC HAMAP; MF_00038; -; 1.
CC InterPro; IPR000715; Glyco_trans_4.
CC InterPro; IPR003524; PNACpPpt_trans.
CC Pfam; PF00953; Glycos_transf_4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; 1.
CC PROSITE; PS01348; MRAY_2; 1.
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
CC Complete proteome.
CC TRANSMEM 3 23 POTENTIAL.
CC TRANSMEM 55 75 POTENTIAL.
CC TRANSMEM 80 100 POTENTIAL.
CC TRANSMEM 117 137 POTENTIAL.
CC TRANSMEM 156 176 POTENTIAL.
CC TRANSMEM 187 207 POTENTIAL.
CC TRANSMEM 231 251 POTENTIAL.
CC TRANSMEM 255 275 POTENTIAL.
CC TRANSMEM 280 300 POTENTIAL.
CC TRANSMEM 334 354 POTENTIAL.
CC SEQUENCE 359 AA; 37713 MW; 27E979BC77BA28C6 CRC64;
Query Match 69.8%; Score 37; DB 1; Length 359;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PFHHH 6
DB 315 PFHHH 319
RESULT 15
NY4R_HUMAN

ID NY4R HUMAN STANDARD; PRT; 375 AA.
AC P50391; Q13456;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuropeptide Y receptor type 4 (NPY4-R) (Pancreatic polypeptide
DE receptor 1) (PPI).
GN PPYR1 OR NPY4R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96070761; PubMed=7592911;
RA Bard J.A., Walker M.W., Branche T.A., Weinshank R.L.;
RT "Cloning and functional expression of a human Y4 subtype receptor for
RT pancreatic polypeptide, neuropeptide Y, and peptide YY.";
RL J. Biol. Chem. 270:26762-26765(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96094298; PubMed=7493937;
RA Lundell I., Blomqvist A.G., Berglund M.M., Schober D.A.,
RA Johnson D., Statnick M.A., Galski R.A., Gehlert D.R., Larhammar D.;
RT "Cloning of a human receptor of the NPY receptor family with high
RT affinity for pancreatic polypeptide and peptide YY.";
RL J. Biol. Chem. 270:29123-29128(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96209788; PubMed=8643460;
RA Yan H., Yang J., Marasco J., Yamaguchi K., Brenner S., Collins F.,
RA Karbon W.;
RT "Cloning and functional expression of cDNAs encoding human and rat
RT pancreatic polypeptide receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4661-4665(1996).
CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE RANK
CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS
CC PP, PP (2-36) AND [ILE-31, GLN-34] PP > [PRO-34] PYY > PYY AND
CC [LEU-31, PRO-34] NPY > NPY > PYY (3-36) AND NPY (2-36) > PP (13-
CC 36) > PP (31-36) > NPY FREE ACID.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN BRAIN, CORONARY ARTERY
CC AND ILEUM. LOW LEVELS IN PANCREAS AND KIDNEY. DETECTED IN COLON
CC AND SMALL INTESTINE.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U35232; AAC50280.1; -.
CC EMBL; Z66526; CAA91433.1; -.
CC EMBL; U42387; AAB07759.1; -.
CC PIR; G02300; G02300.
CC PIR; I39182; I39182.
CC Genew; HGNC:9329; PPYR1.
CC MIM; 601790; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0008015; P:circulation; TAS.
CC GO; GO:0007586; P:digestion; TAS.
CC GO; GO:0007631; P:feeding behavior; TAS.
CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm 1; 1.
CC PRINTS; PR00237; GPCR_RHODOPSIN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 41 63 1 (POTENTIAL).
FT DOMAIN 64 74 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 75 95 2 (POTENTIAL).
FT DOMAIN 96 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 137 3 (POTENTIAL).
FT DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 158 178 4 (POTENTIAL).
FT DOMAIN 179 213 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 214 235 5 (POTENTIAL).
FT DOMAIN 236 265 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 266 288 6 (POTENTIAL).
FT DOMAIN 289 301 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 302 325 7 (POTENTIAL).
FT DOMAIN 326 375 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 2 2 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 19 19 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 114 201 BY SIMILARITY.
FT LIPID 340 340 PALMITATE (POTENTIAL).
FT CONFLICT 99 99 A -> S (IN REF. 3).
SQ SEQUENCE 375 AA; 42195 MW; 0081DBA51436EE22 CRC64;

Query Match 69.8%; Score 37; DB 1; Length 375;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHHR 7
Db 205 WPLAHR 211

Search completed: February 11, 2004, 17:04:13
Job time : 6.16667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	79.2	310	2 T46035	AtMYB84 - Arabidopsis
2	42	79.2	423	2 E64157	hypothetical prote
3	40	75.5	1865	1 I48155	transcription init
4	40	75.5	1893	1 A40262	transcription init
5	39	73.6	316	2 C84009	pantothenate kinas
6	39	73.6	750	2 T04980	hypothetical prote
7	38.5	72.6	750	2 T38435	coiled coil protei
8	38	71.7	284	2 A70976	hypothetical prote
9	38	71.7	293	2 S64925	hypothetical prote
10	38	71.7	316	2 G82204	transcription regu
11	38	71.7	674	2 T21217	hypothetical prote
12	38	71.7	744	2 S65669	biotin sulfoxide r
13	38	71.7	753	2 AH0097	probable biotin su
14	38	71.7	766	2 E70916	probable biotin su
15	38	71.7	925	2 S50490	hypothetical prote
16	37.5	70.8	79	2 T30118	hypothetical prote
17	37	69.8	236	2 T04929	hypothetical prote
18	37	69.8	313	2 F85205	hypothetical prote
19	37	69.8	327	2 F96814	hypothetical prote
20	37	69.8	359	2 F70304	phospho-N-acetylmu
21	37	69.8	359	2 A87023	hypothetical prote
22	37	69.8	359	2 H70579	probable murX prot
23	37	69.8	375	2 G02300	pancreatic polypep
24	37	69.8	375	2 I39182	neuropeptide Y/pep
25	37	69.8	395	2 S39659	protein - Bac
26	37	69.8	520	2 F86401	protein T22C5.11 l
27	37	69.8	557	2 AC2137	ABC transporter AT
28	37	69.8	611	2 G96032	probable thiamine
29	37	69.8	710	2 T47610	hypothetical prote

30	37	69.8	1040	2 T39936	probable helicase
31	37	69.8	1091	2 T48444	hypothetical prote
32	36.5	68.9	332	2 D70716	hypothetical prote
33	36	67.9	101	2 T49585	hemoglobin alpha c
34	36	67.9	243	2 T38519	hypothetical prote
35	36	67.9	297	2 B84088	transcription regu
36	36	67.9	339	2 C95958	probable oxidoredu
37	36	67.9	392	2 C75531	probable N-acyl-L-
38	36	67.9	404	2 T50335	mating pheromone r
39	36	67.9	526	2 S67272	3',5'-cyclic-nucle
40	36	67.9	676	2 A96722	unknown protein T1
41	35	66.0	140	2 I46880	T-cell receptor be
42	35	66.0	254	2 A49169	cholestenone salph
43	35	66.0	300	2 T33232	hypothetical prote
44	35	66.0	319	1 RWRBB	T-cell receptor be
45	35	66.0	369	1 S23248	transcription fact

ALIGNMENTS

RESULT 1

T46035
AtMYB84 - Arabidopsis thaliana
N;Alternate names: protein T16K5.40
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000
C;Accession: T46035
R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke, submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23015
A;Accession: T46035
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-310 <RIE>
A;Cross-references: EMBL:AL132965
A;Experimental source: cultivar Columbia; BAC clone T16K5
C;Genetics:
A;Map position: 3
A;Introns: 46/1; 89/2
A;Note: T16K5.40
C;Superfamily: Arabidopsis myb-related protein Y13; myb DNA-binding repeat homology

Query Match 79.2%; Score 42; DB 2; Length 310;
Best Local Similarity 83.3%; Pred. No. 6.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
DB 163 WPLHHH 168

RESULT 2

E64157
hypothetical protein HI0723 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Sep-1998
C;Accession: E64157
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: E64157
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-423 <TIGR>
A;Cross-references: GB:U32755; GB:L42023; NID:gl573722; PID:gl573726; TIGR:HI0723
A;Note: best homolog was a hypothetical protein from Escherichia coli
C;Genetics:
A;Start codon: GTG

C;Superfamily: potassium uptake protein trkG

Query Match 79.2%; Score 42; DB 2; Length 423;
Best Local Similarity 71.4%; Pred. No. 9.5;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
Db 55 WPCHHK 61

RESULT 3

I48155
transcription initiation factor IID 250K chain splice form 2 - golden hamster
N;Alternate names: 210K nuclear DNA-binding cell cycle gene 1 protein (CCG1); CCG1; TATA
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 15-Oct-1999
A;Accession: I48155
R;Hayashida, T.; Sekiguchi, T.; Noguchi, E.; Sunamoto, H.; Ohba, T.; Nishimoto, T.
Gene 141, 267-270, 1994
A;Title: The CCG1/TAFlII250 gene is mutated in thermosensitive G1 mutants of the BHK21 ce
A;Reference number: I48155; MUID:94215915; PMID:8163200
A;Accession: I48155
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1865 <RES>
A;Cross-references: GB:D26114; NID:G439485; PIDN:BAA05110.1; PID:G474971
C;Superfamily: transcription initiation factor IID 250K chain; bromodomain homology; HMG

C;Keywords: alternative splicing; cell cycle control; DNA binding; duplication; phosphor
F;1190-1269/Domain: HMG box homology <HMG1>
F;1346-1353/Region: nuclear location signal
F;1400-1455/Domain: bromodomain homology <BRO1>
F;1523-1578/Domain: bromodomain homology <BRO2>
F;131,1714,1725,1821,1843/Binding site: phosphate (Ser) (covalent) (by casein kinase II)
F;652,1028,1658/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase
F;994,1335/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status
F;1355,1374/Binding site: phosphate (Thr) (covalent) (by calmodulin-dependent kinase) #s

Query Match 75.5%; Score 40; DB 1; Length 1865;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||
Db 1521 WPFHH 1525

RESULT 4

A40262
transcription initiation factor IID 250K chain splice form 1 - human
N;Alternate names: 210K nuclear DNA-binding cell cycle gene 1 protein (CCG1); TATA-bind
N;Contains: transcription initiation factor IID 250K chain splice form 2
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: A40262; S03005; S00830; S32352; S32353
R;Sekiguchi, T.; Nohiro, Y.; Nakamura, Y.; Hisamoto, N.; Nishimoto, T.
Mol. Cell. Biol. 11, 3317-3325, 1991
A;Title: The human CCG1 gene, essential for progression of the G-1 phase, encodes a 210-
A;Reference number: A40262; MUID:91246200; PMID:2038334
A;Accession: A40262

A;Molecule type: mRNA
A;Residues: 1-177,199-1893 <SEK>
A;Cross-references: GB:D90359; NID:G559319; PIDN:BAA14374.1; PID:G219528
A;Note: nucleotide sequence not complete
R;Sekiguchi, T.; Miyata, T.; Nishimoto, T.
submitted to the EMBL Data Library, February 1988
A;Reference number: S03005
A;Accession: S03005
A;Molecule type: mRNA
A;Residues: 'MYR', 60-177,199-1604,'DNECSKANDIVCLIQYCSSQIEELRF' <SES>
A;Cross-references: EMBL:X07024; NID:G29732; PIDN:CAA30073.1; PID:G29733
A;Note: this sequence has been revised in reference A40262
R;Sekiguchi, T.; Miyata, T.; Nishimoto, T.

EMBO J. 7, 1683-1687, 1988
A;Title: Molecular cloning of the cDNA of human X chromosomal gene (CCG1) which comple
A;Reference number: S00830; MUID:89005056; PMID:3169001
A;Accession: S00830
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1372-1379,1386-1604,'DNE', 'CSSKANDIVCLIQYCSSQIEELRF' <SE6>
A;Cross-references: EMBL:X07024
R;Ruppert, S.; Wang, E.H.; Tjian, R.
Nature 362, 175-179, 1993
A;Title: Cloning and expression of human TAF(II)250: a TBP-associated factor implicated
A;Reference number: S32352; MUID:93196704; PMID:7680771
A;Accession: S32352

A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 178-198 <RUP>
R;Hisatake, K.; Hasegawa, S.; Takada, R.; Nakatani, Y.; Horikoshi, M.; Roeder, R.G.
Nature 362, 179-181, 1993
A;Title: The p250 subunit of native TATA box-binding factor TFIID is the cell-cycle reg
A;Reference number: S32353; MUID:93196705; PMID:8450888
A;Accession: S32353
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'P';587-595;1009-1022;1351-1355;1357-1360 <HIS>
A;Note: 1351-Val, 1353-Lys, and 1354-Glu were also found
C;Genetics:

A;Gene: GDB:TAF2A; CCG1; CCGS; NSCL2; TAFII250
A;Cross-references: GDB:120573; OMIM:313650
A;Map position: Xql3.1-Xql3.1
C;Superfamily: transcription initiation factor IID 250K chain; bromodomain homology; HM
C;Keywords: alternative splicing; cell cycle control; DNA binding; duplication; phospho
F;1-1893/Product: transcription initiation factor IID 250K chain splice form 1 #status
F;1-177,199-1893/Product: transcription initiation factor IID 250K chain splice form 2
F;1216-1295/Domain: HMG box homology <HMG1>
F;1372-1379/Region: nuclear location signal
F;1426-1481/Domain: bromodomain homology <BRO1>
F;1549-1604/Domain: bromodomain homology <BRO2>
F;137,1740,1751,1847,1871/Binding site: phosphate (Ser) (covalent) (by casein kinase II
F;678,1054,1684/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinas
F;1020,1361/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status
F;1381,1400/Binding site: phosphate (Thr) (covalent) (by calmodulin-dependent kinase) #

Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||||
Db 1547 WPFHH 1551

RESULT 5

C84009
pantothenate kinase BH2875 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: C84009
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: C84009

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <STO>
A;Cross-references: GB:AP001516; GB:BA000004; NID:G10175192; PIDN:BAB06594.1; GSPDB:GNO
A;Experimental source: strain C-125
C;Genetics:

A;Gene: BH2875
C;Superfamily: pantothenate kinase

Query Match 73.6%; Score 39; DB 2; Length 316;
Best Local Similarity 83.3%; Pred. No. 22;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
||| |||

Db 115 WPEHHH 120

RESULT 6

T04980

hypothetical protein T16L1.120 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Nov-1999

C;Accession: T04980

R;Bevan, M.; Obermaier, B.; Deutschenbaur, S.; Piravandi, E.; Hobeisel, J.; Mewes, H.W.; submitted to the Protein Sequence Database, November 1998

A;Reference number: Z15393

A;Accession: T04980

A;Molecule type: DNA

A;Residues: 1-750 <BEV>

A;Cross-references: EMBL:AL031394

A;Experimental source: cultivar Columbia; BAC clone T16L1

C;Genetics:

A;Map position: 4

A;Introns: 6/1; 20/2; 181/3; 233/3; 270/3; 306/2; 333/3; 575/3; 617/1; 673/3; 698/3

A;Note: T16L1.120

C;Superfamily: Arabidopsis thaliana hypothetical protein T16L1.120

Query Match 73.6%; Score 39; DB 2; Length 750;

Best Local Similarity 83.3%; Pred. No. 53;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFHHHR 7
|:|:|

Db 21 PYHHHR 26

RESULT 7

T38435

coiled coil protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T38435

R;McDougall, R.; Wood, V.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, December 1997

A;Reference number: Z21793

A;Accession: T38435

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-750 <MCD>

A;Cross-references: EMBL:AL009227; PIDN:CAA15821.1; GSPDB:GN00066; SPDB:SPAC27D7.02C

A;Experimental source: strain 972h-; cosmid c27D7

C;Genetics:

A;Gene: SPDB:SPAC27D7.02C

A;Map position: 1

Query Match 72.6%; Score 38.5; DB 2; Length 750;

Best Local Similarity 60.0%; Pred. No. 65;

Matches 6; Conservative 1; Mismatches 0; Indels 3; Gaps 1;

QY 1 WP---PHHR 7
||| |

Db 308 WPKYVFHHK 317

RESULT 8

A70976

hypothetical protein Rv3435c - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 28-Jul-2000

C;Accession: A70976

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: A70976

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-284 <COL>

A;Cross-references: GB:Z95389; GB:AL123456; NID:G3242261; PIDN:CAB08684.1; PID:G2104356

A;Experimental source: strain H37RV

C;Genetics:

A;Gene: Rv3435c

C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0713

Query Match 71.7%; Score 38; DB 2; Length 284;

Best Local Similarity 71.4%; Pred. No. 29;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHHR 7
||| |

Db 123 WPFDDHR 129

RESULT 9

S64925

hypothetical protein YLR091w - yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein L2525

C;Species: Saccharomyces cerevisiae

C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002

C;Accession: S64925

R;Benes, V.; Rechmann, S.; Nentwich, U.; Schwager, C.; Ansorge, W.; Voss, H. submitted to the Protein Sequence Database, May 1996

A;Reference number: S64920

A;Accession: S64925

A;Molecule type: DNA

A;Residues: 1-293 <BEN>

A;Cross-references: EMBL:Z73263; NID:G1360464; PIDN:CAA97652.1; PID:e245544; PID:G13604

A;Experimental source: strain S288C

C;Genetics:

A;Gene: MIPS:YLR091w

A;Cross-references: SGD:S0004081

A;Map position: 12R

Query Match 71.7%; Score 38; DB 2; Length 293;

Best Local Similarity 71.4%; Pred. No. 30;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHHR 7
||| |

Db 102 WPFENHR 108

RESULT 10

G82204

transcription regulator LysR family VC1390 [imported] - Vibrio cholerae (strain N16961

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: G82204

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: G82204

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-316 <HEI>

A;Cross-references: GB:AE004218; GB:AE003852; NID:g9655881; PIDN:AAF94548.1; GSPDB:GN00

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1390

A;Map position: 1

Query Match 71.7%; Score 38; DB 2; Length 316;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
| | | | |
Db 212 WIFHHH 217

RESULT 11

T21217
hypothetical protein F21G4.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T21217
R/Mortimore, B.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19392
A/Accession: T21217
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-674 <WIL>
A/Cross-references: EMBL:Z81016; PIDN:CAB02665.1; GSPDB:GN00028; CESP:F21G4.1
A/Experimental source: clone F21G4
C/Genetics:
A/Gene: CESP:F21G4.1
A/Map position: X
A/Introns: 58/3; 149/1; 181/3; 213/3; 286/1; 327/2; 457/2; 509/3; 608/1

Query Match 71.7%; Score 38; DB 2; Length 674;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 7
| | | | |
Db 458 WHFYHH 464

RESULT 12

S65669
biotin sulfoxide reductase (EC 1.8.4.-) - Rhodobacter sphaeroides
C/Species: Rhodobacter sphaeroides
C/Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C/Accession: S65669
R/Pollock, V.V.; Barber, M.J.
Arch. Biochem. Biophys. 318, 322-332, 1995
A/Title: Molecular cloning and expression of biotin sulfoxide reductase from Rhodobacter
A/Reference number: S65669; MUID:95251380; PMID:7733660
A/Accession: S65669
A/Molecule type: DNA
A/Residues: 1-744 <POL>
A/Cross-references: EMBL:U08189; NID:9953223; PIDN:AAA74739.1; PID:9953224
A/Experimental source: strain Forma sp. denitrificans
A/Note: the authors translated the initiation codon GTG for residue 1 as Val
C/Genetics:
A/Start codon: GTG
C/Superfamily: trimethylamine-N-oxide reductase
C/Keywords: molybdenum; oxidoreductase

Query Match 71.7%; Score 38; DB 2; Length 744;
Best Local Similarity 83.3%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFFHHH 7
| | | | |
Db 401 PFFHHQ 406

RESULT 13

AH0097
probable biotin sulfoxide reductas2 bisc [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C/Accession: AH0097
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AH0097
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-753 <KUR>
A/Cross-references: GB:AL590842; PIDN:CAC89643.1; PID:gl5978871; GSPDB:GN00175
C/Genetics:
A/Gene: bisc
C/Superfamily: trimethylamine-N-oxide reductase

Query Match 71.7%; Score 38; DB 2; Length 753;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFFHHH 7
| | | | |
Db 425 PFFHHQ 430

RESULT 14

E70916
probable bisc protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C/Accession: E70916
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: E70916
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-766 <COL>
A/Cross-references: GB:Z95844; GB:AL123456; NID:G3250713; PIDN:CAB09264.1; PID:G2131054
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: bisc
C/Superfamily: trimethylamine-N-oxide reductase

Query Match 71.7%; Score 38; DB 2; Length 766;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PFFHHH 7
| | | | |
Db 427 PFFHHQ 432

RESULT 15

S50490
hypothetical protein YER032w - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae
C/Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001
C/Accession: S50490
R/Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A/Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambda
A/Reference number: S50433
A/Accession: S50490
A/Molecule type: DNA
A/Residues: 1-925 <DIE>
A/Cross-references: EMBL:U18778; NID:G603592; PIDN:AAB64565.1; PID:G603624; MIPS:YER032
C/Genetics:

A, Gene: SGD:FIR1
A, Cross-references: SGD:S0000834; MIPS:YER032w
A, Map position: 5R

Query Match	71.7%;	Score 38;	DB 2;	Length 925;
Best Local Similarity	71.4%;	Pred. No. 97;		
Matches	5;	Conservative	1;	Mismatches
			1;	Indels
			0;	Gaps
			0;	

QY 1 WPFHHR 7
| | | | |
Db 179 WKFHNR 185

Search completed: February 11, 2004, 17:11:46
Job time : 13.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-9

Perfect score: 53

Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:**

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	100.0	7	10	US-09-901-187B-9
2	45	84.9	27	12	US-10-029-386-30005
3	42	79.2	309	12	US-10-374-780A-2268
4	42	79.2	309	16	US-10-278-536-120
5	40	75.5	111	12	US-10-209-201C-17
6	40	75.5	111	12	US-10-209-201C-18
7	40	75.5	128	15	US-10-012-542-465
8	40	75.5	165	15	US-10-012-542-464
9	40	75.5	898	15	US-10-043-487-277
10	39	73.6	16	10	US-09-948-018-25
11	38.5	72.6	115	9	US-09-989-722-95
12	38.5	72.6	115	9	US-09-989-723-95
13	38.5	72.6	115	9	US-09-989-279-95
14	38.5	72.6	115	9	US-09-989-727-95
15	38.5	72.6	115	10	US-09-989-731-95

16	38.5	72.6	115	10	US-09-989-732-95	Sequence 95, Appl
17	38.5	72.6	115	10	US-09-991-073-95	Sequence 95, Appl
18	38.5	72.6	115	10	US-09-990-442-95	Sequence 95, Appl
19	38.5	72.6	115	10	US-09-991-163-95	Sequence 95, Appl
20	38.5	72.6	115	10	US-09-993-604-95	Sequence 95, Appl
21	38.5	72.6	115	10	US-09-990-456-95	Sequence 95, Appl
22	38.5	72.6	115	10	US-09-989-721-95	Sequence 95, Appl
23	38.5	72.6	115	10	US-09-992-598-95	Sequence 95, Appl
24	38.5	72.6	115	10	US-09-989-293A-95	Sequence 95, Appl
25	38.5	72.6	115	10	US-09-989-735-95	Sequence 95, Appl
26	38.5	72.6	115	10	US-09-990-444-95	Sequence 95, Appl
27	38.5	72.6	115	10	US-09-991-181-95	Sequence 95, Appl
28	38.5	72.6	115	10	US-09-989-730-95	Sequence 95, Appl
29	38.5	72.6	115	10	US-09-990-436-95	Sequence 95, Appl
30	38.5	72.6	115	10	US-09-993-687-95	Sequence 95, Appl
31	38.5	72.6	115	11	US-09-989-734-95	Sequence 95, Appl
32	38.5	72.6	115	11	US-09-997-653-95	Sequence 95, Appl
33	38.5	72.6	115	11	US-09-993-667-95	Sequence 95, Appl
34	38.5	72.6	115	11	US-09-997-428-95	Sequence 95, Appl
35	38.5	72.6	115	11	US-09-997-666-95	Sequence 95, Appl
36	38.5	72.6	115	11	US-09-990-438-95	Sequence 95, Appl
37	38.5	72.6	115	11	US-09-990-562-95	Sequence 95, Appl
38	38.5	72.6	115	11	US-09-990-711-95	Sequence 95, Appl
39	38.5	72.6	115	11	US-09-989-726-95	Sequence 95, Appl
40	38.5	72.6	115	11	US-09-998-156-95	Sequence 95, Appl
41	38.5	72.6	115	11	US-09-990-437-95	Sequence 95, Appl
42	38.5	72.6	115	11	US-09-991-157-95	Sequence 95, Appl
43	38.5	72.6	115	11	US-09-997-514-95	Sequence 95, Appl
44	38.5	72.6	115	11	US-09-997-573-95	Sequence 95, Appl
45	38.5	72.6	115	11	US-09-991-172-95	Sequence 95, Appl

ALIGNMENTS

RESULT 1
US-09-901-187B-9
; Sequence 9, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Wolozin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-9

Query Match 100.0%; Score 53; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
Db 1 WPFHHR 7

RESULT 2
US-10-029-386-30005
; Sequence 30005, Application US/10029386

Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30005
LENGTH: 27
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ALL36530.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.79
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.6
US-10-029-386-30005

Query Match 84.9%; Score 45; DB 12; Length 27;
Best Local Similarity 83.3%; Pred. No. 3.5;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHHH 6
||:||||
Db 5 WPHYHH 10

RESULT 3
US-10-374-780A-2268
Sequence 2268, Application US/10374780A
Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
APPLICANT: Broun, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09

PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2268
LENGTH: 309
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G234
US-10-374-780A-2268

Query Match 79.2%; Score 42; DB 12; Length 309;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
||:||||
Db 163 WPLHHH 168

RESULT 4
US-10-278-536-120
Sequence 120, Application US/10278536
Publication No. US20030131386A1
GENERAL INFORMATION:
APPLICANT: Samaha, Raymond
APPLICANT: Heard, Jacqueline
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Pineda, Omaira
APPLICANT: Reuber, Lynne
APPLICANT: Riechmann, Jose-Luis
APPLICANT: Yu, Guo-Liang
APPLICANT: Keddie, James
APPLICANT: Ratcliffe, Oliver
APPLICANT: Pilgrim, Marsha
APPLICANT: Adam, Luc
APPLICANT: Broun, Pierre
TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-011
CURRENT APPLICATION NUMBER: US/10/278,536
CURRENT FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 120
LENGTH: 309
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G234
US-10-278-536-120

Query Match 79.2%; Score 42; DB 16; Length 309;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
||:||||
Db 163 WPLHHH 168

RESULT 5
US-10-209-201C-17
Sequence 17, Application US/10209201C
Publication No. US20040009613A1
GENERAL INFORMATION:
APPLICANT: Verdin, Eric
APPLICANT: Bruland, Joan
APPLICANT: Ott, Melanie

```
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-17

Query Match      75.5%; Score 40; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WPFHH 5
Db      29 WPFHH 33

RESULT 6
US-10-209-201C-18
; Sequence 18, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Mesocricetus auratus
US-10-209-201C-18

Query Match      75.5%; Score 40; DB 12; Length 111;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WPFHH 5
Db      29 WPFHH 33

RESULT 7
US-10-012-542-465
; Sequence 465, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-464

Query Match      75.5%; Score 40; DB 15; Length 165;
Best Local Similarity 71.4%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHHR 7
Db      87 WPFHHR 93

RESULT 9
US-10-043-487-277
```

```
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 465
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-465

Query Match      75.5%; Score 40; DB 15; Length 128;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHHR 7
Db      50 WPFHHR 56

RESULT 8
US-10-012-542-464
; Sequence 464, Application US/10012542
; Publication No. US20030044851A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: PZ029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 464
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-542-464

Query Match      75.5%; Score 40; DB 15; Length 165;
Best Local Similarity 71.4%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHHR 7
Db      87 WPFHHR 93

RESULT 9
US-10-043-487-277
```

```

; Sequence 277, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; TITLE OF INVENTION: mammalian polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 277
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Shigella Flexneri
; US-10-043-487-277

Query Match          75.5%; Score 40; DB 15; Length 898;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 WPFHH 5
      |||||
Db      552 WPFHH 556

RESULT 10
US-09-948-018-25
; Sequence 25, Application US/09948018
; Patent No. US20020150977A1
; GENERAL INFORMATION:
; APPLICANT: Theill et al
; TITLE OF INVENTION: TNF RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37677
; CURRENT APPLICATION NUMBER: US/09/948,018
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,191
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-948-018-25

Query Match          73.6%; Score 39; DB 10; Length 16;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 WPFHH 6
      |||||
Db      8 WPNHH 13

RESULT 11
US-09-989-722-95
; Sequence 95, Application US/09989722
; Patent No. US20020072067A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
```

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C63
; CURRENT APPLICATION NUMBER: US/09/989,722
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
```

/	PRIOR APPLICATION NUMBER:	60/088734
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088738
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088742
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088810
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088824
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088826
/	PRIOR FILING DATE:	1998-06-10
/	PRIOR APPLICATION NUMBER:	60/088858
/	PRIOR FILING DATE:	1998-06-11
/	PRIOR APPLICATION NUMBER:	60/088861
/	PRIOR FILING DATE:	1998-06-11
/	PRIOR APPLICATION NUMBER:	60/088876
/	PRIOR FILING DATE:	1998-06-11
/	PRIOR APPLICATION NUMBER:	60/089105
/	PRIOR FILING DATE:	1998-06-12
/	PRIOR APPLICATION NUMBER:	60/089440
/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089512
/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089514
/	PRIOR FILING DATE:	1998-06-16
/	PRIOR APPLICATION NUMBER:	60/089532
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089538
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089598
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089599
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089600
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089653
/	PRIOR FILING DATE:	1998-06-17
/	PRIOR APPLICATION NUMBER:	60/089801
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089907
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089908
/	PRIOR FILING DATE:	1998-06-18
/	PRIOR APPLICATION NUMBER:	60/089947
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089948
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/089952
/	PRIOR FILING DATE:	1998-06-19
/	PRIOR APPLICATION NUMBER:	60/090246
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090252
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090254
/	PRIOR FILING DATE:	1998-06-22
/	PRIOR APPLICATION NUMBER:	60/090349
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090355
/	PRIOR FILING DATE:	1998-06-23
/	PRIOR APPLICATION NUMBER:	60/090429
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090431
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090435
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090444
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090445
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090472
/	PRIOR FILING DATE:	1998-06-24
/	PRIOR APPLICATION NUMBER:	60/090535

```

, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090540
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090542
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090676
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090678
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090690
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090694
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090695
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090696
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090862
, PRIOR FILING DATE: 1998-06-26
, PRIOR APPLICATION NUMBER: 60/090863
, PRIOR FILING DATE: 1998-06-26
, PRIOR APPLICATION NUMBER: 60/091360
, PRIOR FILING DATE: 1998-07-01
, PRIOR APPLICATION NUMBER: 60/091478
, PRIOR FILING DATE: 1998-07-02
, PRIOR APPLICATION NUMBER: 60/091544
, PRIOR FILING DATE: 1998-07-01
, PRIOR APPLICATION NUMBER: 60/091519
, PRIOR FILING DATE: 1998-07-02
, PRIOR APPLICATION NUMBER: 60/091626
, PRIOR FILING DATE: 1998-07-02
, PRIOR APPLICATION NUMBER: 60/091633
, PRIOR FILING DATE: 1998-07-02
, PRIOR APPLICATION NUMBER: 60/091978
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/091982
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/092182
, PRIOR FILING DATE: 1998-07-07

```

Query Match 72.6%; Score 38.5; DB 9; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels

Qy 1 WP-----FHHR 7
Db 59 WPFRRRGHLGI FHHHR 74

RESULT 12

US-09-989-723-95

; Sequence 95, Application US/09989723

; Patent No. US20020072092A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnovers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Geritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: P2730PIC62
 CURRENT APPLICATION NUMBER: US/09/989,723
 CURRENT FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/049787
 PRIOR FILING DATE: 1997-06-16
 PRIOR APPLICATION NUMBER: 60/062250
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/065186
 PRIOR FILING DATE: 1997-11-12
 PRIOR APPLICATION NUMBER: 60/065311
 PRIOR FILING DATE: 1997-11-13
 PRIOR APPLICATION NUMBER: 60/066770
 PRIOR FILING DATE: 1997-11-24
 PRIOR APPLICATION NUMBER: 60/075945
 PRIOR FILING DATE: 1998-02-25
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088021
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088030
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088734
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088742
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090676
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090678
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090690
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090694
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090695
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090696
 ; PRIOR FILING DATE: 1998-06-25
 ; PRIOR APPLICATION NUMBER: 60/090862
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/090863
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: 60/091360
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091478
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091544
 ; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 9; Length 115;

Best Local Similarity 43.8%; Pred. No. 99;

Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 WP-----FHHR 7
 |||||
 Db 59 WPFRRGHLGIFHHR 74

RESULT 13

US-09-989-279-95
 ; Sequence 95, Application US/09989279
 ; Patent No. US20020072496A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE REFERENCE: P2730PIC56
 ; CURRENT APPLICATION NUMBER: US/09/989,279
 ; CURRENT FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088858
 ; PRIOR FILING DATE: 1998-06-11

;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089440
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089600
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089801
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694

;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 9; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;

Matches 7; Conservative 0; Mismatches 9; Indels 0; Gaps 1;

QY 1 WP-----FHHHR 7
|||
Db 59 WPFRRRGLGIFHHHR 74

RESULT 14

US-09-989-727-95
; Sequence 95, Application US/09989727
; Patent No. US20020072497A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16


```

; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      72.6%; Score 38.5; DB 9; Length 115;
Best Local Similarity 43.8%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 WP-----FHHR 7
   || |||||
Db 59 WPFRRRGLGFHHR 74

RESULT 15
US-09-989-731-95
; Sequence 95, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C70
; CURRENT APPLICATION NUMBER: US/09/989,731
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
```

;	PRIOR APPLICATION NUMBER: 60/089538
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089598
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089599
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089600
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089653
;	PRIOR FILING DATE: 1998-06-17
;	PRIOR APPLICATION NUMBER: 60/089801
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089907
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089908
;	PRIOR FILING DATE: 1998-06-18
;	PRIOR APPLICATION NUMBER: 60/089947
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089948
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/089952
;	PRIOR FILING DATE: 1998-06-19
;	PRIOR APPLICATION NUMBER: 60/090246
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090252
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090254
;	PRIOR FILING DATE: 1998-06-22
;	PRIOR APPLICATION NUMBER: 60/090349
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090355
;	PRIOR FILING DATE: 1998-06-23
;	PRIOR APPLICATION NUMBER: 60/090429
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090431
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090435
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090444
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090445
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090472
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090535
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090540
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090542
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090557
;	PRIOR FILING DATE: 1998-06-24
;	PRIOR APPLICATION NUMBER: 60/090676
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090678
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090690
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090694
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090695
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090696
;	PRIOR FILING DATE: 1998-06-25
;	PRIOR APPLICATION NUMBER: 60/090862
;	PRIOR FILING DATE: 1998-06-26
;	PRIOR APPLICATION NUMBER: 60/090863
;	PRIOR FILING DATE: 1998-06-26
;	PRIOR APPLICATION NUMBER: 60/091360
;	PRIOR FILING DATE: 1998-07-01
;	PRIOR APPLICATION NUMBER: 60/091478
;	PRIOR FILING DATE: 1998-07-02
;	PRIOR APPLICATION NUMBER: 60/091544

; PRIOR FILING DATE: 1998-07-01
 ; PRIOR APPLICATION NUMBER: 60/091519
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091626
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091633
 ; PRIOR FILING DATE: 1998-07-02
 ; PRIOR APPLICATION NUMBER: 60/091978
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/091982
 ; PRIOR FILING DATE: 1998-07-07
 ; PRIOR APPLICATION NUMBER: 60/092182
 ; PRIOR FILING DATE: 1998-07-09

Query Match	72.6%	Score 38.5;	DB 10;	Length 115;
Best Local Similarity	43.8%;	Pred. No. 99;		
Matches 7;	Conservative	0;	Mismatches	0;
			Indels	9

Qy 1 WP-----FHHHR 7
Db 59 WPFRRRGLGIFHHHR 74

Search completed: February 11, 2004, 17:54:12
Job time : 25.6667 secs

OM protein - protein search, using sw model
Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	84.9	580	4	US-09-252-991A-22036
2	40	75.5	65	1	US-08-227-536-4
3	40	75.5	65	5	PCT-US95-04682-4
4	40	75.5	128	4	US-09-461-325-465
5	40	75.5	165	4	US-09-461-325-464
6	40	75.5	1872	1	US-08-188-582-14
7	40	75.5	1872	1	US-08-646-715-14
8	40	75.5	1893	1	US-08-188-582-11
9	40	75.5	1893	1	US-08-646-715-11
10	38.5	72.6	115	4	US-09-461-325-160
11	38.5	72.6	115	4	US-09-996-243-95
12	37	69.8	350	2	US-08-495-695B-33
13	37	69.8	375	1	US-08-176-412-2
14	37	69.8	375	2	US-08-555-268A-2
15	37	69.8	375	2	US-08-495-695B-2
16	37	69.8	375	4	US-09-200-673-17
17	37	69.8	375	5	PCT-US94-14436-2
18	37	69.8	470	4	US-09-252-991A-28702
19	36	67.9	16	3	US-09-177-249-113
20	36	67.9	287	4	US-09-690-454-55
21	36	67.9	417	4	US-09-252-991A-20165
22	36	67.9	1209	4	US-09-252-991A-25844
23	35.5	67.0	106	4	US-09-634-238-349
24	35	66.0	128	3	US-08-905-223-306
25	35	66.0	200	4	US-09-252-991A-31738
26	35	66.0	219	4	US-09-252-991A-32844
27	35	66.0	241	3	US-08-634-475-7

28	35	66.0	241	4	US-09-709-791-7	Sequence 7, Appli
29	35	66.0	254	1	US-07-795-859B-6	Sequence 6, Appli
30	35	66.0	254	1	US-08-457-616-6	Sequence 6, Appli
31	35	66.0	262	4	US-09-252-991A-26036	Sequence 26036, A
32	35	66.0	368	4	US-09-252-991A-20452	Sequence 20452, A
33	35	66.0	427	4	US-09-252-991A-32813	Sequence 32813, A
34	35	66.0	481	4	US-09-252-991A-24508	Sequence 24508, A
35	35	66.0	525	4	US-09-252-991A-29635	Sequence 29635, A
36	35	66.0	694	4	US-09-252-991A-30587	Sequence 30587, A
37	35	66.0	919	2	US-08-788-674-4	Sequence 4, Appli
38	35	66.0	942	4	US-09-171-461-12	Sequence 12, Appli
39	35	66.0	951	3	US-08-816-346-58	Sequence 58, Appli
40	35	66.0	951	3	US-09-335-411-58	Sequence 58, Appli
41	35	66.0	952	2	US-08-788-674-5	Sequence 5, Appli
42	35	66.0	952	3	US-08-816-346-4	Sequence 4, Appli
43	35	66.0	952	3	US-09-335-411-4	Sequence 4, Appli
44	35	66.0	967	3	US-08-816-346-56	Sequence 56, Appli
45	35	66.0	967	3	US-09-335-411-56	Sequence 56, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-22036
; Sequence 22036, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22036
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22036

Query Match 84.9%; Score 45; DB 4; Length 580;
Best Local Similarity 85.7%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
Db 294 WPAHHR 300

RESULT 2
US-08-227-536-4
; Sequence 4, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-227-536-4

Query Match 75.5%; Score 40; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 4 WPFHH 8

RESULT 4

US-09-461-325-465
Sequence 465, Application US/09461325A
Patent No. 6475753
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/09/461,325A
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER APPLICATION NUMBER: 60/089,507
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,508
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,509
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,510
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/090,112
EARLIER FILING DATE: 1998-06-22
EARLIER APPLICATION NUMBER: 60/090,113
EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 465
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
US-09-461-325-465

Query Match 75.5%; Score 40; DB 4; Length 128;
Best Local Similarity 71.4%; Pred. No. 8.1;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
Db 50 WPFHHR 56

RESULT 5

US-09-461-325-464
Sequence 464, Application US/09461325A
Patent No. 6475753
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/09/461,325A
CURRENT FILING DATE: 1999-12-14
EARLIER APPLICATION NUMBER: PCT/US99/13418
EARLIER FILING DATE: 1999-06-15
EARLIER APPLICATION NUMBER: 60/089,507
EARLIER FILING DATE: 1998-06-16
EARLIER APPLICATION NUMBER: 60/089,508

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Williams Ph.D., Kathleen A.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: DFCI-308XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-227-536-4

Query Match 75.5%; Score 40; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 4 WPFHH 8

RESULT 3

PCT-US95-04682-4
Sequence 4, Application PC/TUS9504682
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
TITLE OF INVENTION: FACTOR P300 AND USES OF P300
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,536
FILING DATE: 14-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Holliday C. Heine, Ph.D.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-308Xq999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 65 amino acids

EARLIER FILING DATE: 1998-06-16
 EARLIER APPLICATION NUMBER: 60/089,509
 EARLIER FILING DATE: 1998-06-16
 EARLIER APPLICATION NUMBER: 60/089,510
 EARLIER FILING DATE: 1998-06-16
 EARLIER APPLICATION NUMBER: 60/090,112
 EARLIER FILING DATE: 1998-06-22
 EARLIER APPLICATION NUMBER: 60/090,113
 EARLIER FILING DATE: 1998-06-22
 NUMBER OF SEQ ID NOS: 532
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 464
 LENGTH: 165
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-461-325-464

Query Match 75.5%; Score 40; DB 4; Length 165;
 Best Local Similarity 71.4%; Pred. No. 10;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7
 ||:|
 Db 87 WPHRHR 93

RESULT 6
 US-08-188-582-14
 Sequence 14, Application US/08188582
 Patent No. 5534410
 GENERAL INFORMATION:
 APPLICANT: Tjian, Robert
 APPLICANT: Comai, Lucio
 APPLICANT: Dynlacht, Brian D.
 APPLICANT: Hoey, Timothy
 APPLICANT: Ruppert, Siegfried
 APPLICANT: Tanese, Naoko
 APPLICANT: Wang, Edith
 APPLICANT: Weinzierl, Robert O.J.
 TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/188,582
 FILING DATE: 28-JAN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1872 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide

US-08-188-582-14
 Query Match 75.5%; Score 40; DB 1; Length 1872;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
 |||||
 Db 1526 WPFHH 1530

RESULT 7
 US-08-646-715-14
 Sequence 14, Application US/08646715
 Patent No. 5637686
 GENERAL INFORMATION:
 APPLICANT: Tjian, Robert
 APPLICANT: Comai, Lucio
 APPLICANT: Dynlacht, Brian D.
 APPLICANT: Hoey, Timothy
 APPLICANT: Ruppert, Siegfried
 APPLICANT: Tanese, Naoko
 APPLICANT: Wang, Edith
 APPLICANT: Weinzierl, Robert O.J.
 TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
 TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
 STREET: 4 Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94111-4187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,715
 FILING DATE: 09-MAY-1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/188,582
 FILING DATE: 28-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
 TELEPHONE: (415) 781-1989
 TELEFAX: (415) 398-3249
 TELEX: 910 277299
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1872 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-646-715-14

Query Match 75.5%; Score 40; DB 1; Length 1872;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
 |||||
 Db 1526 WPFHH 1530

RESULT 8

US-08-188-582-11
; Sequence 11, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1893 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-582-11

Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||
Db 1547 WPFHH 1551

RESULT 9
US-08-646-715-11
; Sequence 11, Application US/08646715
; Patent No. 5637686
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1893 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-715-11

Query Match 75.5%; Score 40; DB 1; Length 1893;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
|||
Db 1547 WPFHH 1551

RESULT 10
US-09-461-325-160
; Sequence 160, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens

```
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-461-325-160

Query Match          72.6%; Score 38.5; DB 4; Length 115;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 WP-----FHHHR 7
   |||
Db 59 WPFRRRGHLGFHHHR 74

RESULT 11
US-09-996-243-95
; Sequence 95, Application US/09996243
; Patent No. 6478825
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C13
; CURRENT APPLICATION NUMBER: US/09/996,243
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
```

;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089908
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089948
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090254
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match 72.6%; Score 38.5; DB 4; Length 115;
Best Local Similarity 43.8%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 WP-----FHHHR 7
|||
Db 59 WPFRRRGLGIFHHHR 74

RESULT 12
US-08-495-695B-33
; Sequence 33, Application US/08495695B
; Patent No. 5976814
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; APPLICANT: Walker, Mary
; APPLICANT: Branchek, Theresa
; APPLICANT: Weinshank, Richard L.
; TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
; TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
; TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,695B
; FILING DATE: 13-Jan-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44743-A-PCT-US/JPW/JHB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-495-695B-33

Query Match 69.8%; Score 37; DB 2; Length 350;
Best Local Similarity 71.4%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHHR 7
|||
Db 205 WPLAHR 211

RESULT 13
US-08-176-412-2
; Sequence 2, Application US/08176412
; Patent No. 5516653
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; APPLICANT: Walker, Mary
; APPLICANT: Branchek, Theresa

APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,412
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743/JPW/TEP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-176-412-2

Query Match 69.8%; Score 37; DB 1; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
Db 205 WPLAHR 211

RESULT 14
US-08-555-268A-2
Sequence 2, Application US/08555268A
Patent No. 5958709
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
APPLICANT: Walker, Mary
APPLICANT: Branchek, Theresa
APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/555,268A

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743-Z/JPW/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-555-268A-2

Query Match 69.8%; Score 37; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
|||
Db 205 WPLAHR 211

RESULT 15
US-08-495-695B-2
Sequence 2, Application US/08495695B
Patent No. 5976814
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
APPLICANT: Walker, Mary
APPLICANT: Branchek, Theresa
APPLICANT: Weinshank, Richard L.
TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY/PANCREATIC POLYPEPTIDE
TITLE OF INVENTION: RECEPTOR (Y4) AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/495,695B
FILING DATE: 13-Jan-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44743-A-PCT-US/JPW/JHB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-495-695B-2

Query Match 69.8%; Score 37; DB 2; Length 375;
Best Local Similarity 71.4%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WPFHHR 7
||
Db 205 WPLAHR 211

Search completed: February 11, 2004, 17:13:38
Job time : 11.4167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-9
Perfect score: 53
Sequence: 1 WPFHHR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	100.0	7	AAE14554	Human alpha-synuclein
2	42	79.2	151	AAG36521	Arabidopsis thalia
3	42	79.2	176	AAG20314	Arabidopsis thalia
4	42	79.2	176	AAG38740	Arabidopsis thalia
5	42	79.2	181	AAG20313	Arabidopsis thalia
6	42	79.2	181	AAG38739	Arabidopsis thalia
7	42	79.2	309	AAB67830	Amino acid sequenc
8	42	79.2	310	AAG20312	Arabidopsis thalia
9	42	79.2	310	AAG38738	Arabidopsis thalia

10	42	79.2	541	22	ABB68621	Drosophila melanog
11	40	75.5	123	22	AAO04423	Human polypeptide
12	40	75.5	128	21	AAV86527	Human gene 72-enco
13	40	75.5	165	21	AAV86526	Human gene 72-enco
14	40	75.5	898	23	ABG70103	Human prey protein
15	40	75.5	1575	23	AAU84357	Protein TAF2A diff
16	40	75.5	1872	15	AAV56493	TATA-binding prote
17	40	75.5	1872	17	AAW06078	Drosophila TATA-bi
18	40	75.5	1872	18	AAW25030	TATA-binding prote
19	40	75.5	1886	22	ABG06009	Novel human diagno
20	40	75.5	1893	15	AAV56491	TATA-binding prote
21	40	75.5	1893	17	AAW06082	Human TATA-binding
22	40	75.5	1893	18	AAW25020	TATA-binding prote
23	40	75.5	1924	22	ABG06008	Novel human diagno
24	39	73.6	11	23	AAU97110	pAMG21-human MK61
25	38.5	72.6	115	21	AAV86243	Human secreted pro
26	38.5	72.6	115	21	AAV66649	Membrane-bound pro
27	38.5	72.6	115	22	AAU29066	Human PRO polypept
28	38.5	72.6	115	22	AAV65172	Human PRO537 (UNQ3
29	38.5	72.6	115	24	ABU71154	Human PRO537 prote
30	38.5	72.6	115	24	ABU65611	Human secreted/tra
31	38.5	72.6	115	24	ABU65944	Novel human secret
32	38.5	72.6	115	24	ABU67448	Human secreted/tra
33	38.5	72.6	115	24	ABU65306	Human PRO polypept
34	38.5	72.6	115	24	ABU59065	Novel human secret
35	38.5	72.6	115	24	ABU59212	Human secreted/tra
36	38.5	72.6	115	24	ABU59361	Novel human secret
37	38.5	72.6	115	24	ABU60496	Human secreted/tra
38	38.5	72.6	115	24	ABU57987	Human PRO polypept
39	38.5	72.6	115	24	ABU58442	Human PRO polypept
40	38.5	72.6	115	24	ABU58918	Human secreted/tr
41	38.5	72.6	115	24	ABU55978	Human secreted/tra
42	38.5	72.6	115	24	ABU56973	Human PRO polypept
43	38.5	72.6	115	24	ABU13878	Human PRO537 polyp
44	38.5	72.6	115	24	ABU10552	Human secreted/tra
45	38.5	72.6	115	24	ABU10833	Human PRO polypept

ALIGNMENTS

RESULT 1
AAE14554

ID AAE14554 standard; peptide; 7 AA.

AC AAE14554;

XX 17-MAY-2002 (first entry)

DT Human alpha-synuclein aggregation inhibitor #9.

DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.

XX Homo sapiens

OS WO200204482-A1.

XX 17-JAN-2002.

PD 06-JUL-2001; 2001WO-US21379.

PF 07-JUL-2000; 2000US-217319P.

PR 28-MAR-2001; 2001US-279199P.

XX (PANA-) PANACEA PHARM INC.

PI Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

PT Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
PS Claim 40; Page 37; 52pp; English.
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 53; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WPFHHHR 7
|||
Db 1 WPFHHHR 7

RESULT 2
AAG36521
ID AAG36521 standard; Protein; 151 AA.

XX AAG36521;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 44768.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 151;
Best Local Similarity 83.3%; Pred. No. 8.9;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
|||
Db 4 WPLHHH 9

RESULT 3
AAG20314
ID AAG20314 standard; Protein; 176 AA.

XX AC AAG20314;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 22452.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.
XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 176;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 29 WPLHHH 34

RESULT 4
AAG38740
ID AAG38740 standard; Protein; 176 AA.
XX
AC AAG38740;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47837.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR	06-AUG-1999;	99US-0147416.	QY	1 WPFHHH 6	
PR	09-AUG-1999;	99US-0147493.	Db	29 WPLHHH 34	
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	13-AUG-1999;	99US-0148684.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149175.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			
PR	31-AUG-1999;	99US-0151438.			
PR	01-SEP-1999;	99US-0151930.			
PR	07-SEP-1999;	99US-0152363.			
PR	10-SEP-1999;	99US-0153070.			
PR	13-SEP-1999;	99US-0153758.			
PR	15-SEP-1999;	99US-0154018.			
PR	16-SEP-1999;	99US-0154039.			
PR	20-SEP-1999;	99US-0154779.			
PR	22-SEP-1999;	99US-0155139.			
PR	23-SEP-1999;	99US-0155486.			
PR	24-SEP-1999;	99US-0155659.			
PR	28-SEP-1999;	99US-0156458.			
PR	29-SEP-1999;	99US-0156596.			
PR	04-OCT-1999;	99US-0157117.			
PR	05-OCT-1999;	99US-0157753.			
PR	06-OCT-1999;	99US-0157865.			
PR	07-OCT-1999;	99US-0158029.			
PR	08-OCT-1999;	99US-0158232.			
PR	12-OCT-1999;	99US-0158369.			
PR	13-OCT-1999;	99US-0159293.			
PR	13-OCT-1999;	99US-0159294.			
PR	13-OCT-1999;	99US-0159295.			
PR	14-OCT-1999;	99US-0159329.			
PR	14-OCT-1999;	99US-0159330.			
PR	14-OCT-1999;	99US-0159331.			
PR	14-OCT-1999;	99US-0159637.			
PR	14-OCT-1999;	99US-0159638.			
PR	18-OCT-1999;	99US-0159584.			
PR	21-OCT-1999;	99US-0160741.			
PR	21-OCT-1999;	99US-0160767.			
PR	21-OCT-1999;	99US-0160768.			
PR	21-OCT-1999;	99US-0160770.			
PR	21-OCT-1999;	99US-0160814.			
PR	21-OCT-1999;	99US-0160815.			
PR	22-OCT-1999;	99US-0160980.			
PR	22-OCT-1999;	99US-0160981.			
PR	22-OCT-1999;	99US-0160989.			
PR	25-OCT-1999;	99US-0161404.			
PR	25-OCT-1999;	99US-0161405.			
PR	25-OCT-1999;	99US-0161406.			
PR	26-OCT-1999;	99US-0161359.			
PR	26-OCT-1999;	99US-0161360.			
PR	26-OCT-1999;	99US-0161361.			
PR	28-OCT-1999;	99US-0161920.			
PR	28-OCT-1999;	99US-0161992.			
PR	28-OCT-1999;	99US-0161993.			
PR	29-OCT-1999;	99US-0162142.			

Query Match 79.2%; Score 42; DB 21; Length 176;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 181;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 WPFHHH 6
Db 34 WPLHHH 39
RESULT 6

AAG38739
ID AAG38739 standard; Protein; 181 AA.
XX
AC AAG38739;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 47836.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 181;
Best Local Similarity 83.3%; Pred. No. 11;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 34 WPLHHH 39

RESULT 7
AAB67830
ID AAB67830 standard; Protein; 309 AA.
XX
AC AAB67830;
XX
DT 29-JUN-2001 (first entry)
XX

DE Amino acid sequence of a plant transcription factor G234.
XX Transcription factor; flowering time; transgenic plant; vernalisation;
KW plant development; plant physiology; flowering.
XX Arabidopsis thaliana.
OS
XX WO200126459-A2.
PN
XX 19-APR-2001.
PD
XX 11-OCT-2000; 2000WO-US28141.
PF
XX 12-OCT-1999; 99US-0159464.
PR 08-NOV-1999; 99US-0164132.
PR 17-NOV-1999; 99US-0166228.
PR 17-APR-2000; 2000US-0197899.
PR 22-AUG-2000; 2000US-0227439.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA (RATC/) RATCLIFFE O.
PA (HEAR/) HEARD J.
PA (SAMA/) SAMAHA R.
PA (CREE/) CREELMAN R.
PA (KEDD/) KEDDIE J.
PA (JIAN/) JIANG C.
PA (REUB/) REUBER L.
PA (RIEC/) RIECHMANN J L.
XX Ratcliffe O, Heard J, Samaha R, Creelman R, Keddie J, Jiang C;
PI Reuber L, Riechmann JL;
XX WPI; 2001-266398/27.
DR N-PSDB; AAF80400.
XX
PT New transgenic plant comprises a recombinant polynucleotide encoding a
PT plant transcription factor polypeptide and has a modified flowering
PT time or vernalization requirement -
XX Example 7; Page 73-75; 108pp; English.
PS
XX The present sequence represents a plant transcription factor protein
CC which modifies the flowering time of a plant. The polynucleotide
CC sequence is used to produce transgenic plants which have a modified
CC flowering time or a modified vernalisation requirement. The
CC polynucleotides and polypeptides are useful for modifying plant
CC development, physiology or biochemistry such that the modified plants
CC have a trait advantage over wild type plants. In particular they are
CC useful for accelerating, delaying or preventing flowering. The
CC polynucleotides are also useful as nucleic acid probes and primers.
CC They may be used to identify proteins that can modify the activity of
CC the transcription factor.
XX
SQ Sequence 309 AA;

Query Match 79.2%; Score 42; DB 22; Length 309;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 8
AAG20312
ID AAG20312 standard; Protein; 310 AA.
XX
AC AAG20312;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22450.

XX	Protein identification; signal transduction pathway; metabolic pathway;	23-JUN-1999;	99US-0140353.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	23-JUN-1999;	99US-0140354.
KW	termination sequence.	24-JUN-1999;	99US-0140695.
XX	Arabidopsis thaliana.	28-JUN-1999;	99US-0140823.
XX	EP1033405-A2.	29-JUN-1999;	99US-0140991.
XX	06-SEP-2000.	30-JUN-1999;	99US-0141287.
PD	25-FEB-2000; 2000EP-0301439.	01-JUL-1999;	99US-0141842.
XX	25-FEB-1999;	01-JUL-1999;	99US-0142154.
PR	05-MAR-1999;	02-JUL-1999;	99US-0142055.
PR	09-MAR-1999;	06-JUL-1999;	99US-0142390.
PR	23-MAR-1999;	08-JUL-1999;	99US-0142803.
PR	25-MAR-1999;	09-JUL-1999;	99US-0142920.
PR	29-MAR-1999;	12-JUL-1999;	99US-0142977.
PR	01-APR-1999;	13-JUL-1999;	99US-0143542.
PR	06-APR-1999;	14-JUL-1999;	99US-0143624.
PR	08-APR-1999;	15-JUL-1999;	99US-0144005.
PR	16-APR-1999;	16-JUL-1999;	99US-0144085.
PR	19-APR-1999;	16-JUL-1999;	99US-0144086.
PR	21-APR-1999;	19-JUL-1999;	99US-0144325.
PR	23-APR-1999;	19-JUL-1999;	99US-0144331.
PR	28-APR-1999;	19-JUL-1999;	99US-0144332.
PR	30-APR-1999;	19-JUL-1999;	99US-0144333.
PR	30-APR-1999;	19-JUL-1999;	99US-0144334.
PR	04-MAY-1999;	19-JUL-1999;	99US-0144335.
PR	05-MAY-1999;	20-JUL-1999;	99US-0144352.
PR	06-MAY-1999;	20-JUL-1999;	99US-0144632.
PR	06-MAY-1999;	20-JUL-1999;	99US-0144884.
PR	11-MAY-1999;	21-JUL-1999;	99US-0144814.
PR	14-MAY-1999;	21-JUL-1999;	99US-0145086.
PR	14-MAY-1999;	21-JUL-1999;	99US-0145088.
PR	14-MAY-1999;	22-JUL-1999;	99US-0145085.
PR	18-MAY-1999;	22-JUL-1999;	99US-0145087.
PR	19-MAY-1999;	22-JUL-1999;	99US-0145089.
PR	20-MAY-1999;	22-JUL-1999;	99US-0145192.
PR	21-MAY-1999;	23-JUL-1999;	99US-0145145.
PR	24-MAY-1999;	23-JUL-1999;	99US-0145218.
PR	25-MAY-1999;	23-JUL-1999;	99US-0145224.
PR	27-MAY-1999;	26-JUL-1999;	99US-0145276.
PR	28-MAY-1999;	27-JUL-1999;	99US-0145913.
PR	21-MAY-1999;	27-JUL-1999;	99US-0145918.
PR	24-MAY-1999;	27-JUL-1999;	99US-0145919.
PR	25-MAY-1999;	28-JUL-1999;	99US-0145951.
PR	27-MAY-1999;	02-AUG-1999;	99US-0146386.
PR	28-MAY-1999;	02-AUG-1999;	99US-0146388.
PR	01-JUN-1999;	02-AUG-1999;	99US-0146389.
PR	03-JUN-1999;	03-AUG-1999;	99US-0147038.
PR	04-JUN-1999;	04-AUG-1999;	99US-0147204.
PR	07-JUN-1999;	05-AUG-1999;	99US-0147302.
PR	08-JUN-1999;	05-AUG-1999;	99US-0147192.
PR	10-JUN-1999;	06-AUG-1999;	99US-0147260.
PR	14-JUN-1999;	06-AUG-1999;	99US-0147303.
PR	16-JUN-1999;	09-AUG-1999;	99US-0147416.
PR	17-JUN-1999;	09-AUG-1999;	99US-0147493.
PR	18-JUN-1999;	10-AUG-1999;	99US-0147935.
PR	18-JUN-1999;	11-AUG-1999;	99US-0148171.
PR	18-JUN-1999;	12-AUG-1999;	99US-0148319.
PR	18-JUN-1999;	13-AUG-1999;	99US-0148341.
PR	18-JUN-1999;	13-AUG-1999;	99US-0148565.
PR	18-JUN-1999;	16-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	27-AUG-1999;	99US-0151080.
PR		30-AUG-1999;	99US-0151303.

3 X

PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		
PF	25-FEB-2000; 2000EP-0301439.	
XX		
PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.

PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 79.2%; Score 42; DB 21; Length 310;
Best Local Similarity 83.3%; Pred.No. 19;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WPFHHH 6
Db 163 WPLHHH 168

RESULT 10
ABB68621
ID ABB68621 standard; Protein; 541 AA.
XX
AC ABB68621;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32655.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL12724.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX Disclosure; SEQ ID NO 32655; 2lpp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 541 AA;

Query Match 79.2%; Score 42; DB 22; Length 541;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
|||
Db 300 WPLHHH 305

RESULT 11
AAO04423
ID AAO04423 standard; Protein; 123 AA.
XX
AC AAO04423;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 18315.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX WO200164835-A2.
PN
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US04927.
PF
XX 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR N-PSDB; AAI84354.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -

XX Claim 20; SEQ ID NO 18315; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 123 AA;
Query Match 75.5%; Score 40; DB 22; Length 123;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHH 6
|||
Db 57 WPFHHH 62

RESULT 12
AAI86527
ID AAI86527 standard; Protein; 128 AA.
XX
AC AAI86527;
XX
DT 19-APR-2000 (first entry)
XX
DE Human gene 72-encoded protein fragment, SEQ ID NO:442.

XX Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW therapy.

XX Homo sapiens.

OS
XX WO9966041-A1.
PN
XX 23-DEC-1999.
PD
XX 15-JUN-1999; 99WO-US13418.
PF
XX 16-JUN-1998; 98US-0089507.
PR 16-JUN-1998; 98US-0089508.
PR 16-JUN-1998; 98US-0089509.
PR 16-JUN-1998; 98US-0089510.
PR 22-JUN-1998; 98US-0090112.
PR 22-JUN-1998; 98US-0090113.

XX (HUMA-) HUMAN GENOME SCI INC.

PA
XX Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;
PI Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R;
PI Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G;
XX WPI; 2000-106100/09.

XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -

XX PS Disclosure; Page 152; 586pp; English.

XX CC AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.

CC CC AAY86215 to AAY86333 are the secreted proteins encoded by the 94 human

CC CC genes. This sequence represents a fragment of one of the human secreted

CC CC proteins. The genes and their corresponding secreted polypeptides are

CC CC useful for preventing, treating or ameliorating medical conditions,

CC CC e.g., by protein or gene therapy. Also pathological conditions can be

CC CC diagnosed by determining the amount of the new polypeptides in a sample

CC CC or by determining the presence of mutations in the new genes. Specific

CC CC uses are described for each of the 94 genes, based on which tissues they

CC CC are most highly expressed in, and include developing products for the

CC CC diagnosis or treatment of cancer, tumours, developmental abnormalities

CC CC and foetal deficiencies, blood disorders, diseases of the immune system,

CC CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive

CC CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin

CC CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney

CC CC disorders, digestive/endocrine disorders, infections and AIDS. The

CC CC polypeptides are also useful for identifying their binding partners.

CC CC The sequences shown in AAY86334 to AAY86585 represent fragments of the

CC CC secreted proteins.

XX XX Sequence 128 AA;

Query Match 75.5%; Score 40; DB 21; Length 128;

Best Local Similarity 71.4%; Pred. No. 16;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7

Db 50 WPTHHR 56

RESULT 13

AAAY86526

ID . AAY86526 standard; Protein; 165 AA.

XX AC AAY86526;

XX DT 19-APR-2000 (first entry)

XX DE Human gene 72-encoded protein fragment, SEQ ID NO:441.

XX KW Human; secreted protein; cancer; tumour; developmental abnormality;

KW foetal deficiency; blood disorder; immune system disorder; inflammation;

KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;

KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;

KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;

KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;

KW therapy.

XX OS Homo sapiens.

XX WO9966041-A1.

XX PD 23-DEC-1999.

XX PF 15-JUN-1999; 99WO-US13418.

XX PR 16-JUN-1998; 98US-0089507.

PR 16-JUN-1998; 98US-0089508.

PR 16-JUN-1998; 98US-0089509.

PR 16-JUN-1998; 98US-0089510.

PR 22-JUN-1998; 98US-0090112.

PR 22-JUN-1998; 98US-0090113.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Ni J, Rosen CA, Wei Y, Young PE, Florence KA;

XX Soppet DR, Brewer LA, Endress GA, Carter KC, Mucenski M, Ebner R;

PI Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G;

XX

DR WPI; 2000-106100/09.

XX New isolated human genes and the secreted polypeptides they encode,

PT useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders

XX Disclosure; Page 152; 586pp; English.

XX AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.

CC AAY86215 to AAY86333 are the secreted proteins encoded by the 94 human

CC CC genes. This sequence represents a fragment of one of the human secreted

CC CC proteins. The genes and their corresponding secreted polypeptides are

CC CC useful for preventing, treating or ameliorating medical conditions,

CC CC e.g., by protein or gene therapy. Also pathological conditions can be

CC CC diagnosed by determining the amount of the new polypeptides in a sample

CC CC or by determining the presence of mutations in the new genes. Specific

CC CC uses are described for each of the 94 genes, based on which tissues they

CC CC are most highly expressed in, and include developing products for the

CC CC diagnosis or treatment of cancer, tumours, developmental abnormalities

CC CC and foetal deficiencies, blood disorders, diseases of the immune system,

CC CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive

CC CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin

CC CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney

CC CC disorders, digestive/endocrine disorders, infections and AIDS. The

CC CC polypeptides are also useful for identifying their binding partners.

CC CC The sequences shown in AAY86334 to AAY86585 represent fragments of the

CC CC secreted proteins.

XX XX Sequence 165 AA;

Query Match 75.5%; Score 40; DB 21; Length 165;

Best Local Similarity 71.4%; Pred. No. 21;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WPFHHR 7

Db 87 WPTHHR 93

RESULT 14

ABG70103

ID ABG70103 standard; Protein; 898 AA.

XX AC ABG70103;

XX DT 21-OCT-2002 (first entry)

XX DE Human prey protein for Shigella ospC1 #21.

XX KW Prey protein; ospB; ospD1; ipaD; ipaC; ipaH9.8; ospG;

KW ospC1; Shigella; shigellosis; bacillary dysentery; antibacterial;

KW yeast two-hybrid system; protein-protein interaction; SID;

KW selected interacting domain; human.

XX OS Homo sapiens.

XX WO200257303-A2.

XX PD 25-JUL-2002.

XX PF 11-JAN-2002; 2002WO-EP00777.

XX PR 12-JAN-2001; 2001US-261130P.

XX PA (HYBR-) HYBRIGENICS.

XX PI Legrain P;

XX WPI; 2002-599706/64.

DR N-PSDB; ABS51496.

XX New complex of protein-protein interactions between a bait Shigella

PT flexneri polypeptide and a prey mammalian or human placenta polypeptide

PT for treating or preventing bacillary dysentery in a mammal or human -
XX Claim 7; Page 80-81; 162pp; English.
PS
XX The invention relates to a complex of protein-protein interactions
CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaD, ipaC,
CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the
CC specification. The complexes are formed using the yeast two-hybrid
CC system. Also included are (1) a recombinant host cell expressing the
CC interactions between the Shigella flexneri polypeptide and a mammalian
CC polypeptide defined in the specification; (2) selecting a modulating
CC compound that inhibits or activates the protein-protein interactions;
CC (3) a modulating compound obtained from the method of (2); (4) a SID
CC (selected interacting domain) polypeptide or its fragment or variant
CC comprising the human polypeptides appearing as ABG70042-ABG70242;
CC (5) a SID polynucleotide or its fragment or variant comprising
CC encoding the above polypeptides a vector comprising (5);
CC (6) a recombinant host cell containing the vector; and (10) a protein
CC chip comprising Shigella flexneri polypeptide and a mammalian polypeptide
CC defined in the specification. A pharmaceutical composition comprising the
CC compound, polypeptide or polynucleotide is useful for treating or
CC preventing shigellosis (bacillary dysentery) in a human or mammal.
CC The present sequence represents a human prey protein isolated by the
CC yeast two-hybrid assay, forming a complex of the invention with a
CC shigella protein.
XX
SQ Sequence 898 AA;

Query Match 75.5%; Score 40; DB 23; Length 898;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 552 WPFHH 556

RESULT 15
AAU84357
ID AAU84357 standard; Protein; 1575 AA.
XX
AC AAU84357;
XX
DT 08-MAY-2002 (first entry)
XX
DE Protein TAP2A differentially expressed in breast cancer tissue.
XX
KW Human; diagnosis of breast cancer; endometrial cancer; breast tumour;
KW MAI; mitotic activity index; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200210436-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-US23642.
XX
PR 28-JUL-2000; 2000US-222093P.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
PA (BAAK/) BAAK J.
XX
PI Baak J, Mutter GL;
XX
DR WPI; 2002-180084/23.
DR N-PSDB; ABK35577.
XX
PT Diagnosing breast cancer comprises determining expression of nucleic
PT acid molecules or expression products that are differentially expressed
PT in normal and malignant tissue -
XX
PS Claim 37; Page 210-215; 219pp; English.

XX The present invention relates to a method for diagnosing breast cancer
CC in a subject suspected of having endometrial cancer. The method
CC comprises determining the expression of a set of human genes or
CC expression products in an endometrial sample suspected of being
CC cancerous. The human genes of the invention are differentially
CC expressed in breast tumours characterised as high or low MAI (mitotic
CC activity index). These sets of genes can be used to discriminate between
CC high and low MAI breast tumours. The invention also provides DNA and
CC protein microarrays for analysing the expression of the human genes and
CC their protein products. The methods and arrays are useful for the
CC diagnosis and prognosis of endometrial cancer, selecting and monitoring
CC treatment regimes, and identification of compounds useful for the
CC treatment of endometrial cancer. AAU84311-AAU84361 represent the human
CC proteins of the invention that are differentially expressed in breast
CC cancer tissue.
XX
SQ Sequence 1575 AA;

Query Match 75.5%; Score 40; DB 23; Length 1575;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WPFHH 5
Db 1491 WPFHH 1495

Search completed: February 11, 2004, 17:03:02
Job time : 42.25 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHHKT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	7	23	AAE14555
2	39	86.7	98	21	AAG56024
3	37	82.2	83	22	AAU57344
4	37	82.2	140	22	ABB65911
5	37	82.2	442	22	ABB62210
6	36	80.0	91	23	ABB53926
7	36	80.0	110	17	AAW02631
8	36	80.0	136	24	AAE34013
9	36	80.0	136	24	AAE34023

10	36	80.0	140	23	AAU52548	Maize ASR protein.
11	36	80.0	210	23	AAU87835	T. aureum 7091 par
12	36	80.0	272	23	AAU87836	T. aureum 7091 elo
13	36	80.0	272	23	AAU87837	T. aureum 7091 elo
14	36	80.0	272	23	AAU87839	T. aureum 7091 elo
15	36	80.0	272	23	AAU87841	T. aureum 7091 elo
16	36	80.0	272	23	AAU87842	T. aureum 7091 elo
17	36	80.0	283	21	AAV79250	Mammalian putative
18	36	80.0	286	9	AAU82590	Polypeptide with g
19	36	80.0	299	23	AAU87834	Mouse elongase MEL
20	36	80.0	750	24	ABU07373	Human protein NOV5
21	36	80.0	873	24	ABU11589	Human MDDT polypep
22	36	80.0	1036	24	ABU11770	Human MDDT polypep
23	35	77.8	59	21	AAU41234	Human ORFX ORF998
24	35	77.8	76	23	ABP34316	Human ORF3289 prot
25	35	77.8	113	22	ABG11726	Novel human diagno
26	35	77.8	119	23	ABP34926	Human ORF3899 prot
27	35	77.8	176	21	AAG42058	Arabidopsis thalia
28	35	77.8	187	21	AAG42057	Arabidopsis thalia
29	35	77.8	206	23	ABG79615	Sunflower ribonuc
30	35	77.8	242	11	AAU06483	18-2-3-/TRF59. Mu
31	35	77.8	277	22	ABB61428	Drosophila melanog
32	35	77.8	292	21	AAG42056	Arabidopsis thalia
33	35	77.8	294	22	ABG24169	Novel human diagno
34	35	77.8	311	22	AAG66563	Human carboxypepti
35	35	77.8	349	17	AAW06182	Modified HCPB (D25
36	35	77.8	349	17	AAW06173	mature HCPB-(His)6
37	35	77.8	349	17	AAW06181	Modified HCPB (D25
38	35	77.8	349	18	AAW13750	Carboxypeptidase B
39	35	77.8	349	18	AAW13751	Carboxypeptidase B
40	35	77.8	349	18	AAW13752	Carboxypeptidase B
41	35	77.8	349	18	AAW13753	Carboxypeptidase B
42	35	77.8	349	18	AAW13757	Carboxypeptidase B
43	35	77.8	349	18	AAW13758	Carboxypeptidase B
44	35	77.8	349	18	AAW13759	Carboxypeptidase B
45	35	77.8	349	18	AAW13760	Carboxypeptidase B

ALIGNMENTS

RESULT 1

AAE14555
ID AAE14555 standard; peptide; 7 AA.
XX AAE14555;
AC AAE14555;
XX 17-MAY-2002 (first entry)
XX Human alpha-synuclein aggregation inhibitor #10.
DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.
XX Hemo sapiens.
OS WO200204482-A1.
PN 17-JAN-2002.
XX 06-JUL-2001; 2001WO-US21379.
XX 07-JUL-2000; 2000US-217319P.
PR 28-MAR-2001; 2001US-279199P.
XX (PANA-) PANACEA PHARM INC.
XX Wolozin B, Ostretova-Golts N, Lebowitz MS;
XX WPI; 2002-179695/23.
XX Determination of an agent capable of inhibiting aggregation of alpha

PT

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
PS Claim 40; Page 37; 52pp; English.
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 45; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 1 HLYHHKT 7

RESULT 2
AAG56024
ID AAG56024 standard; Protein; 98 AA.

XX AC AAG56024;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 71943.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.7%; Score 39; DB 21; Length 98;
Best Local Similarity 71.4%; Pred. No. 6.5;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|:|:|
Db 84 HIYHHQT 90

RESULT 3

AAU57344
ID AAU57344 standard; Protein; 83 AA.

XX AC AAU57344;

XX DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #18240.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.
DR N-PSDB; AAS59582.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -

XX Example 1; SEQ ID No 18539; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 83 AA;

Query Match 82.2%; Score 37; DB 22; Length 83;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HLYHHK 6
|:|:|:|:
Db 49 HLYHHR 54

RESULT 4
ABB65911
ID ABB65911 standard; Protein; 140 AA.

XX
AC ABB65911;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 24525.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL10014.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 24525; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 140 AA;

Query Match 82.2%; Score 37; DB 22; Length 140;
Best Local Similarity 71.4%; Pred. No. 22;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 HLYHHKT 7
|:|:|:|:
Db 37 HVYHHST 43

RESULT 5
ABB62210

ID ABB62210 standard; Protein; 442 AA.

XX
AC ABB62210;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 13422.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL06313.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 13422; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 442 AA;

Query Match 82.2%; Score 37; DB 22; Length 442;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HLYHHKT 7
|:|:|:|:
Db 320 HVYHHST 326

RESULT 6
ABB53926

ID ABB53926 standard; Protein; 91 AA.

XX
AC ABB53926;

XX

DT 16-MAY-2002 (first entry)
 DE Lactococcus lactis protein ygaE.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
 XX
 OS Lactococcus lactis IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR WPI; 2002-043418/06.
 XX
 XX New nucleotide sequence useful in the identification of Lactococcus
 PT lactis and related species -
 PT
 XX
 PS Claim 6; SEQ ID No 628; 2504pp; French.
 XX
 CC The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 91 AA;
 Query Match 80.0%; Score 36; DB 23; Length 91;
 Best Local Similarity 85.7%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HLYHHKT 7
 Db 56 HLCCHKT 62
 RESULT 7
 AAW02631
 ID AAW02631 standard; Protein; 110 AA.
 XX
 AC AAW02631;
 XX
 DT 07-NOV-1996 (first entry)
 XX
 DE Tomato P119 protein.
 XX
 KW P119 gene; promoter; fruit-specific expression; transgenic plant;
 KW herbicide resistance; disease resistance; crop protection; tomato;
 KW tobacco.
 XX
 OS Lycopersicon esculentum cv. Golden Nugget.
 XX
 PN WO9619103-A1.
 XX
 PD 27-JUN-1996.
 XX
 PF 29-NOV-1995; 95WO-US15482.
 XX
 PR 20-DEC-1994; 94US-0359696.

XX (DNAP) DNA PLANT TECHNOLOGY CORP.
 PA Dunsmuir P, Stott JS;
 PI
 XX WPI; 1996-309200/31.
 DR N-PSDB; AAT32863.
 XX
 PT New plant P119 promoter - useful for generating transgenic plant,
 PT pref. tomato or tobacco, with e.g. herbicide, fungal disease or
 PT bacterial disease resistance
 XX
 PS Claim 17; Page 41; 50pp; English.
 XX
 CC P119 (AAW02631) is the product of a novel gene (see also AAT32863)
 CC isolated from the pericarp of cherry tomato cv. Golden Nugget.
 CC P119 mRNA is present at high levels in tomato pericarp (both
 CC mature and at various stages of ripening), at moderate levels in
 CC flowers and immature green fruit, at low levels in roots and stems,
 CC and is not present in leaves. The promoter (see also AAT32864) of
 CC the P119 gene can be used to express operably linked DNA
 CC sequences primarily in fruit tissue of transgenic plants, esp.
 CC tomato and tobacco.
 XX
 SQ Sequence 110 AA;
 Query Match 80.0%; Score 36; DB 17; Length 110;
 Best Local Similarity 83.3%; Pred. No. 26;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HLYHHK 6
 Db 10 HLFHHK 15
 RESULT 8
 AAE34013
 ID AAE34013 standard; Protein; 136 AA.
 XX
 AC AAE34013;
 XX
 DT 02-MAY-2003 (first entry)
 XX
 DE Lolium perenne LpASR protein.
 XX
 KW Abscisic acid-inducible and stress responsive protein; ASR; A22; PKABA;
 KW stress-inducible cysteine protease; late embryogenesis abundant protein;
 KW LEA; dehydrin; DHN; abscisic acid-induced protein kinase; gene therapy;
 KW CYS; seed development; plant tolerance; germination; plant protectant;
 KW ryegrass.
 XX
 OS Lolium perenne.
 XX
 PN WO200290547-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 07-MAY-2002; 2002WO-AU00564.
 XX
 PR 07-MAY-2001; 2001AU-0004821.
 XX
 PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
 PA (AGRE-) AGRESEARCH LTD.
 XX
 PI Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
 XX
 DR WPI; 2003-129183/12.
 DR N-PSDB; AAD52414.
 XX
 PT New isolated nucleic acid encoding ASR, A22, CYS, LEA, DHN or PKABA
 PT proteins, useful as molecular genetic markers, and in modifying plant
 PT and/or seed development and responses to stresses and adverse
 PT environmental stimuli -

XX PS XX
Claim 24; Fig 2; 231pp; English.
The invention relates to nucleic acid encoding abscisic acid-inducible
and stress responsive proteins (ASR and A22), stress-inducible cysteine
proteases (CYS), late embryogenesis abundant proteins (LEA), dehydrins
(DHN) and abscisic acid-induced protein kinases (PKABA). The invention
also relates to a method for modification of plant and seed development
as molecular genetic markers. The method is useful for modifying plant
response to an environmental stimulus, modifying plant tolerance to
abiotic, osmotic and/or temperature stresses, modifying seed dormancy
and/or germination, development, maturation, and modifying a plant
developmental process. They are also useful for modifying plant
tolerance and adaptation to stresses and adverse environmental stimuli.
The invention is also used in gene therapy. The preset sequence is
Lolium perenne LpASR protein.

XX SQ Sequence 136 AA;
Query Match 80.0%; Score 36; DB 24; Length 136;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
Db 8 HLFHHK 13

RESULT 9
AAE34023
ID AAE34023 standard; Protein; 136 AA.
XX
AC AAE34023;
XX
DT 02-MAY-2003 (first entry)
XX
DE Lolium perenne ASRa protein.
XX
KW Abscisic acid-inducible and stress responsive protein; ASR; A22; PKABA;
stress-inducible cysteine protease; late embryogenesis abundant protein;
LEA; dehydrin; DHN; abscisic acid-induced protein kinase; gene therapy;
CYS; seed development; plant tolerance; germination; plant protectant;
ryegrass.
XX
OS Lolium perenne.
XX
PN WO200290547-A1.
XX
PD 14-NOV-2002.
XX
PF 07-MAY-2002; 2003WO-AU00564.
XX
PR 07-MAY-2001; 2001AU-0004821.
XX
PA (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.
PA (AGRE-) AGRESEARCH LTD.
XX
PI Spangenberg G, Sawbridge TI, Ong EK, Emmerling M;
XX
XX WPI; 2003-129183/12.
DR N-PSDB; AAD52469.
XX
XX New isolated nucleic acid encoding ASR, A22, CYS, LEA, DHN or PKABA
proteins, useful as molecular genetic markers, and in modifying plant
and/or seed development and responses to stresses and adverse
environmental stimuli -
XX
PS Claim 24; Fig 35; 231pp; English.
XX
XX The invention relates to nucleic acid encoding abscisic acid-inducible
and stress responsive proteins (ASR and A22), stress-inducible cysteine
proteases (CYS), late embryogenesis abundant proteins (LEA), dehydrins

CC (DHN) and abscisic acid-induced protein kinases (PKABA). The invention
also relates to a method for modification of plant and seed development
and plant responses to stresses and stimuli. The invention is useful
as molecular genetic markers. The method is useful for modifying plant
response to an environmental stimulus, modifying plant tolerance to
abiotic, osmotic and/or temperature stresses, modifying seed dormancy
and/or germination, development, maturation, and modifying a plant
developmental process. They are also useful for modifying plant
tolerance and adaptation to stresses and adverse environmental stimuli.
The invention is also used in gene therapy. The present sequence is
Lolium perenne ASRa protein.

XX SQ Sequence 136 AA;
Query Match 80.0%; Score 36; DB 24; Length 136;
Best Local Similarity 83.3%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
Db 8 HLFHHK 13

RESULT 10
AAM52548
ID AAM52548 standard; Protein; 140 AA.
XX
AC AAM52548;
XX
DT 01-FEB-2002 (first entry)
XX
DE Maize ASR protein.
XX
KW Maize; ASR; abscisic acid-water stress-ripening-induced protein;
water stress resistance; crop plant.
XX
OS Zea mays.
XX
PN WO200183756-A1.
XX
PD 08-NOV-2001.
XX
PF 24-APR-2001; 2001WO-FR01252.
XX
PR 28-APR-2000; 2000FR-0005534.
XX
PA (BIOG-) BIOGEMMA.
XX
PI Zivvy M, Perez P;
XX
DR WPI; 2002-041496/05.
DR N-PSDB; ABA01409.
XX
PT Plant with improved resistance to water stress, contains sense or
antisense sequence encoding abscisic acid-water stress-ripening-induced
protein -
XX
PS Claim 3; Page 32; 42pp; French.
XX
XX The present invention relates to a method for the preparation of a plant
containing an altered amount of ASR protein (abscisic acid-water
stress-ripening-induced protein). Plants with altered expression of ASR
have increased resistance to water stress, relative to a non-transformed
plant. The method is especially applied to crop plants such as maize
(especially), wheat, rape, sunflower and peas to increase resistance to
water stress. The present sequence is a protein sequence for maize ASR
which was used to illustrate the present invention.

XX SQ Sequence 140 AA;
Query Match 80.0%; Score 36; DB 23; Length 140;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
Db 10 HLFHHK 15

RESULT 11
AAU87835
ID AAU87835 standard; Protein; 210 AA.
XX
AC AAU87835;
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 partial elongase.
XX
KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
OS Thraustochytrium aureum.
XX
PN WO200208401-A2.
XX
PD 31-JAN-2002.
XX
PF 24-JUL-2001; 2001WO-US23259.
XX
PR 24-JUL-2000; 2000US-0624670.
PR 11-JUL-2001; 2001US-0903456.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
DR WPI; 2002-172011/22.
DR N-PSDB; ABK46404.
XX
PT Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
PS Example 21; Fig 66; 27lpp; English.
XX
CC The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated
CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase
CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.

SQ Sequence 210 AA;
Query Match 80.0%; Score 36; DB 23; Length 210;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 167 HVYHHAT 173

RESULT 12
AAU87836
ID AAU87836 standard; Protein; 272 AA.
XX
AC AAU87836;
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 elongase TELO1 from plasmid PRAT-4-A1.
XX
KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
OS Thraustochytrium aureum.
XX
PN WO200208401-A2.
XX
PD 31-JAN-2002.
XX
PF 24-JUL-2001; 2001WO-US23259.
XX
PR 24-JUL-2000; 2000US-0624670.
PR 11-JUL-2001; 2001US-0903456.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
DR WPI; 2002-172011/22.
DR N-PSDB; ABK46405.
XX
PT Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
PS Example 21; Fig 74; 27lpp; English.
XX
CC The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated
CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase
CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,

CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX Sequence 272 AA;
SQ
Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 152 HVYHHAT 158
RESULT 13
AAU87837
ID AAU87837 standard; Protein; 272 AA.
XX
AC AAU87837;
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 elongase TELO1 from plasmid PRAT-4-A2.
XX
KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
OS Thraustochytrium aureum.
XX
PN WO200208401-A2.
XX
PD 31-JAN-2002.
XX
PF 24-JUL-2001; 2001WO-US23259.
XX
PR 24-JUL-2000; 2000US-0624670.
PR 11-JUL-2001; 2001US-0903456.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
DR WPI; 2002-172011/22.
DR N-PSDB; ABK46406.
XX
PT Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
PS Example 21; Fig 75; 271pp; English.
XX
CC The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated
CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase
CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa), for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of

CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX Sequence 272 AA;
SQ
Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 152 HVYHHAT 158
RESULT 14
AAU87839
ID AAU87839 standard; Protein; 272 AA.
XX
AC AAU87839;
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 elongase TELO1 from plasmid PRAT-4-A4.
XX
KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX
OS Thraustochytrium aureum.
XX
PN WO200208401-A2.
XX
PD 31-JAN-2002.
XX
PF 24-JUL-2001; 2001WO-US23259.
XX
PR 24-JUL-2000; 2000US-0624670.
PR 11-JUL-2001; 2001US-0903456.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
DR WPI; 2002-172011/22.
DR N-PSDB; ABK46408.
XX
PT Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
PS Example 21; Fig 77; 271pp; English.
XX
CC The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated
CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase

CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa); for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX
SQ Sequence 272 AA;

Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

RESULT 15
AAU87841
ID AAU87841 standard; Protein; 272 AA.
XX AAU87841;
XX
DT 05-JUN-2002 (first entry)
XX
DE T. aureum 7091 elongase TEL01 from plasmid pRAT-4-A6.
XX
KW Elongase; polyunsaturated fatty acid; pufa; transgenic plant;
KW transgenic non-human animal; plant oil; arachidonic acid; nutritional;
KW pharmaceutical; cosmetic; animal feed; restenosis; angioplasty; AIDS;
KW acquired immunodeficiency syndrome; multiple sclerosis; cachexia; asthma;
KW inflammatory skin disease; osteoporosis; kidney stone; cancer; eczema;
KW inflammation; rheumatoid arthritis; psoriasis; platelet aggregation;
KW vasodilation; cholesterol reduction; gastrointestinal bleeding.
XX

OS Thraustochytrium aureum.
XX
XX WO200208401-A2.
XX
PD 31-JAN-2002.
XX
PF 24-JUL-2001; 2001WO-US23259.
XX
PR 24-JUL-2000; 2000US-0624670.
PR 11-JUL-2001; 2001US-0903456.
XX
PA (ABBO) ABBOTT LAB.
XX
XX Mukerji P, Das T, Huang Y, Parker-Barnes JM, Leonard AE;
PI Thurmond J, Pereira SL;
XX
DR WPI; 2002-172011/22.
DR N-PSDB; ABK46371.

XX Isolated nucleic acid sequences encoding elongase proteins, useful in
PT the production of polyunsaturated fatty acids, e.g. arachidonic acid,
PT which can then be used nutritional compositions and pharmaceutical
PT compositions (for treating AIDS) -
XX
PS Example 21; Fig 79; 271pp; English.
XX
CC The invention relates to isolated nucleic acid sequences encoding
CC elongase proteins and the elongase proteins themselves. Also
CC included are a purified polypeptide which elongates polyunsaturated

CC fatty acids and has at least 30% amino acid similarity to the amino acid
CC sequence to an elongase protein, a vector comprising the nucleic
CC acid, a plant cell (or tissue or whole plant) comprising the vector and
CC expressing the nucleic acid, a plant oil or acid expressed by the
CC transgenic plant, and a transgenic non-human animal expressing elongase
CC in its fluid. The nucleic acids and elongase proteins may be used
CC directly or indirectly in the production of polyunsaturated fatty acids
CC (pufa); for e.g. arachidonic acid, which can then be used nutritional
CC compositions, pharmaceutical compositions, cosmetics and animal feeds.
CC The pharmaceutical compositions may be used in the treatment of
CC restenosis after angioplasty, AIDS (acquired immunodeficiency syndrome),
CC multiple sclerosis, inflammatory skin diseases, osteoporosis, kidney or
CC urinary tract stones, malignant cancer, cachexia associated with cancer,
CC eczema, symptoms of inflammation, rheumatoid arthritis, asthma and
CC psoriasis. They are also useful for inhibiting platelet aggregation,
CC inducing vasodilation, reducing cholesterol, reducing or preventing
CC gastrointestinal bleeding and side effects of non-steroidal
CC anti-inflammatory drugs. The present sequence represents an elongase
CC of the invention.
XX
SQ Sequence 272 AA;

Query Match 80.0%; Score 36; DB 23; Length 272;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

Search completed: February 11, 2004, 17:03:03
Job time : 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHHK 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	36	80.0	110	1	US-08-359-696-2
2	36	80.0	283	4	US-09-145-828A-22
3	35	77.8	349	4	US-09-011-769A-47
4	35	77.8	349	4	US-09-011-769A-60
5	35	77.8	349	4	US-09-011-769A-64
6	34	75.6	272	4	US-09-145-828A-18
7	33	73.3	458	5	PCT-US96-00994-4
8	33	73.3	855	4	US-08-890-865A-10
9	33	73.3	913	4	US-08-971-089-4
10	32	71.1	157	4	US-09-252-991A-27738
11	32	71.1	237	4	US-09-252-991A-18630
12	32	71.1	272	4	US-09-177-419C-2
13	32	71.1	298	4	US-09-177-419C-4
14	32	71.1	402	4	US-09-252-991A-26501
15	32	71.1	481	4	US-09-252-991A-23317
16	32	71.1	517	4	US-09-134-001C-3188
17	32	71.1	616	4	US-09-252-991A-19164
18	32	71.1	969	2	US-08-284-941-2
19	32	71.1	969	2	US-08-447-642-2
20	32	71.1	969	3	US-09-236-503-2
21	32	71.1	969	5	PCT-US93-02147A-2
22	31	68.9	147	4	US-09-145-828A-20
23	31	68.9	159	4	US-09-205-258-270
24	31	68.9	327	4	US-09-247-155-94
25	31	68.9	327	4	US-09-599-360B-12
26	31	68.9	348	4	US-09-252-991A-31001
27	31	68.9	354	4	US-09-107-532A-4236

28	31	68.9	427	4	US-09-328-352-5312	Sequence 5312, Ap
29	31	68.9	465	4	US-09-252-991A-30946	Sequence 30946, A
30	31	68.9	527	4	US-09-134-001C-3358	Sequence 3358, Ap
31	31	68.9	589	4	US-09-252-991A-30001	Sequence 30001, A
32	31	68.9	731	4	US-09-708-426-12	Sequence 12, Appl
33	31	68.9	787	4	US-09-252-991A-28939	Sequence 28939, A
34	30	66.7	72	3	US-09-042-071-38	Sequence 38, Appl
35	30	66.7	88	4	US-08-311-731A-293	Sequence 293, Ap
36	30	66.7	103	4	US-09-328-352-4929	Sequence 4929, Ap
37	30	66.7	115	4	US-09-996-243-95	Sequence 95, Appl
38	30	66.7	124	3	US-08-881-450A-2	Sequence 2, Appli
39	30	66.7	124	3	US-09-042-071-39	Sequence 39, Appl
40	30	66.7	163	4	US-09-328-352-6165	Sequence 6165, Ap
41	30	66.7	174	4	US-09-544-716-20	Sequence 20, Appl
42	30	66.7	178	3	US-09-042-071-42	Sequence 42, Appl
43	30	66.7	195	3	US-09-042-071-40	Sequence 40, Appl
44	30	66.7	207	3	US-09-013-881-4	Sequence 4, Appli
45	30	66.7	207	4	US-09-612-473-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-359-696-2
; Sequence 2, Application US/08359696
; Patent No. 5633440
; GENERAL INFORMATION:
; APPLICANT: DUNSMUIR, Pamela
; APPLICANT: STOTT, Jamie S.
; TITLE OF INVENTION: P-119 PROMOTERS AND THEIR USES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/359,696
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 12176-004800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-359-696-2

Query Match 80.0%; Score 36; DB 1; Length 110;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHK 6
||:||||
Db 10 HLFHHK 15
RESULT 2

US-09-145-828A-22
; Sequence 22, Application US/09145828A
; Patent No. 6403349
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda E. Y.
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Thurmond, Jennifer
; APPLICANT: Kirchner, Stephen J.
; APPLICANT: Parker-Barnes, Jennifer M.
; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF
; FILE REFERENCE: 6407.US.01
; CURRENT APPLICATION NUMBER: US/09/145,828A
; CURRENT FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (282)...(282)
; OTHER INFORMATION: Xaa = Unknown or other at position 282
US-09-145-828A-22

Query Match 80.0%; Score 36; DB 4; Length 283;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
|:|||||
Db 125 HVYHAT 131

RESULT 3
US-09-011-769A-47
; Sequence 47, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; BLAKEY, David C.
; DAVIES, David H.
; HENNAM, John F.
; HENNEQUIN, Laurent F.A.
; MARSHAM, Peter R.
; DOWELL, Robert I.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996

; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-011-769A-47

Query Match 77.8%; Score 35; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
|:|||||
Db 327 HLYHH 331

RESULT 4
US-09-011-769A-60
; Sequence 60, Application US/09011769A
; Patent No. 6436691
; GENERAL INFORMATION:
; APPLICANT: SLATER, Anthony M.
; BLAKEY, David C.
; DAVIES, David H.
; HENNAM, John F.
; HENNEQUIN, Laurent F.A.
; MARSHAM, Peter R.
; DOWELL, Robert I.
; TITLE OF INVENTION: Chemical Compounds
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, LLP
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,769A
; FILING DATE: 13-Feb-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/01975
; FILING DATE: 13-AUG-1996
; APPLICATION NUMBER: GB 9612295.7
; FILING DATE: 12-JUN-1996
; APPLICATION NUMBER: GB 9611019.2
; FILING DATE: 25-MAY-1996
; APPLICATION NUMBER: GB 9516810.0
; FILING DATE: 16-AUG-1995
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-011-769A-60

Query Match 77.8%; Score 35; DB 4; Length 349;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5

Db 327 HLYHH 331

RESULT 5

US-09-011-769A-64

; Sequence 64, Application US/09011769A

; Patent No. 6436691

; GENERAL INFORMATION:

; APPLICANT: SLATER, Anthony M.

; BLAKEY, David C.

; DAVIES, David H.

; HENNAM, John F.

; HENNEQUIN, Laurent F.A.

; MARSHAM, Peter R.

; DOWELL, Robert I.

; TITLE OF INVENTION: Chemical Compounds

; NUMBER OF SEQUENCES: 87

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pillsbury Madison & Sutro, LLP

; STREET: 1100 New York Ave., N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 Mb disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MS Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/011,769A

; FILING DATE: 13-Feb-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB96/01975

; FILING DATE: 13-AUG-1996

; APPLICATION NUMBER: GB 9612295.7

; FILING DATE: 12-JUN-1996

; APPLICATION NUMBER: GB 9611019.2

; FILING DATE: 25-MAY-1996

; APPLICATION NUMBER: GB 9516810.0

; FILING DATE: 16-AUG-1995

; INFORMATION FOR SEQ ID NO: 64:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 349 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 64:

US-09-011-769A-64

Query Match 77.8%; Score 35; DB 4; Length 349;

Best Local Similarity 100.0%; Pred. No. 60;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HLYHH 5

Db 327 HLYHH 331

RESULT 6

US-09-145-828A-18

; Sequence 18, Application US/09145828A

; Patent No. 6403349

; GENERAL INFORMATION:

; APPLICANT: Abbott Laboratories

; APPLICANT: Mukerji, Pradip

; APPLICANT: Leonard, Amanda E. Y.

; APPLICANT: Huang, Yung-Sheng

; APPLICANT: Thurmond, Jennifer

; APPLICANT: Kirchner, Stephen J.

; APPLICANT: Parker-Barnes, Jennifer M.

; TITLE OF INVENTION: THE ELONGASE GENE AND USES THEREOF

; FILE REFERENCE: 6407.US.01

; CURRENT APPLICATION NUMBER: US/09/145,828A

; CURRENT FILING DATE: 1998-09-02

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 18

; LENGTH: 272

; TYPE: PRT

; ORGANISM: Murine

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (272)...(272)

; OTHER INFORMATION: Xaa = Unknown or other at position 272

US-09-145-828A-18

Query Match 75.6%; Score 34; DB 4; Length 272;

Best Local Similarity 71.4%; Pred. No. 70;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HLYHHKT 7

Db 146 HWYHST 152

RESULT 7

PCT-US96-00994-4

; Sequence 4, Application PC/TUS9600994

; Sequence 4, Application US/

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: SELF-RENEWING PLURIPOTENT HEMATOPOIETIC

; TITLE OF INVENTION: STEM CELL COMPOSITIONS, METHODS OF USE,

; TITLE OF INVENTION: AND CULTURE SYSTEMS THEREFOR

; NUMBER OF SEQUENCES: 4

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/00994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US SN 08/462,108

; FILING DATE: 05-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US SN 08/378,144

; FILING DATE: 24-JAN-1995

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 458 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; PCT-US96-00994-4

Query Match 73.3%; Score 33; DB 5; Length 458;

Best Local Similarity 80.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HLYHH 5

Db 160 HWYHH 164

RESULT 8

US-08-890-865A-10

; Sequence 10, Application US/08890865A

; Patent No. 6307019

; GENERAL INFORMATION:

; APPLICANT: Constantini, Franklin

; APPLICANT: Zeng, Li

; TITLE OF INVENTION: AXIN GENE AND USES THEREOF

; NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: US
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/890,865A
 FILING DATE: 10-JUL-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/54249
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)278-0400
 TELEFAX: (212)391-0526
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 855 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 US-08-890-865A-10

Query Match 73.3%; Score 33; DB 4; Length 855;
 Best Local Similarity 83.3%; Pred. No. 3.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 :|||
 Db 549 NLYHHK 554

RESULT 9
 US-08-971-089-4
 ; Sequence 4, Application US/08971089
 ; Patent No. 6376174
 ; GENERAL INFORMATION:
 ; APPLICANT: Pulst, Stefan M.
 ; APPLICANT: Scoles, Daniel R.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING
 ; TITLE OF INVENTION: SCHWANNOMIN-BINDING-PROTEINS AND PRODUCTS RELATED THERETO
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CAMPBELL & FLORES, LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: United States
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/971,089
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/030,987
 ; FILING DATE: 15-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ramos, Robert T.
 ; REGISTRATION NUMBER: 37,915

REFERENCE/DOCKET NUMBER: P-CE 2862
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619)535-9001
 TELEFAX: (619)535-8949
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 913 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-971-089-4

Query Match 73.3%; Score 33; DB 4; Length 913;
 Best Local Similarity 80.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
 :|||
 Db 572 HIYHH 576

RESULT 10
 US-09-252-991A-27738
 ; Sequence 27738, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 27738
 ; LENGTH: 157
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-27738

Query Match 71.1%; Score 32; DB 4; Length 157;
 Best Local Similarity 66.7%; Pred. No. 89;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
 :|||
 Db 37 HLYHHK 42

RESULT 11
 US-09-252-991A-18630
 ; Sequence 18630, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 18630
 ; LENGTH: 237
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-18630

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26501
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26501

Query Match 71.1%; Score 32; DB 4; Length 402;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||:
Db 300 HLLHHR 305

RESULT 15

US-09-252-991A-23317
; Sequence 23317, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23317
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23317

Query Match 71.1%; Score 32; DB 4; Length 481;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||:
Db 74 HLLHHR 79

Search completed: February 11, 2004, 17:13:39
Job time : 11.4167 secs

Query Match 71.1%; Score 32; DB 4; Length 237;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||:
Db 213 HLLHHR 218

RESULT 12

US-09-177-419C-2
; Sequence 2, Application US/09177419C
; Patent No. 6562609
; GENERAL INFORMATION:
; APPLICANT: Russel, David W
; APPLICANT: Lund, Erik G
; TITLE OF INVENTION: Cholesterol 25-Hydroxylase
; FILE REFERENCE: UTSD1370
; CURRENT APPLICATION NUMBER: US/09/177,419C
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 272
; TYPE: PRT
; ORGANISM: human
US-09-177-419C-2

Query Match 71.1%; Score 32; DB 4; Length 272;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||:
Db 143 HLLHHR 148

RESULT 13

US-09-177-419C-4
; Sequence 4, Application US/09177419C
; Patent No. 6562609
; GENERAL INFORMATION:
; APPLICANT: Russel, David W
; APPLICANT: Lund, Erik G
; TITLE OF INVENTION: Cholesterol 25-Hydroxylase
; FILE REFERENCE: UTSD1370
; CURRENT APPLICATION NUMBER: US/09/177,419C
; CURRENT FILING DATE: 1998-10-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Mouse
US-09-177-419C-4

Query Match 71.1%; Score 32; DB 4; Length 298;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||:
Db 143 HLLHHR 148

RESULT 14

US-09-252-991A-26501
; Sequence 26501, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

OM protein - protein search, using sw model

Run on: February 11, 2004, 17:09:56 ; Search time 24.6667 Seconds
(without alignments)
59.419 Million cell updates/sec

Title: US-09-901-187C-10

Perfect score: 45

Sequence: 1 HLYHHKT 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09E_PUBCOMB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	100.0	7	10	US-09-901-187B-10
2	36	80.0	210	10	US-09-903-456-68
3	36	80.0	210	12	US-10-156-911-68
4	36	80.0	272	10	US-09-903-456-75
5	36	80.0	272	10	US-09-903-456-76
6	36	80.0	272	10	US-09-903-456-78
7	36	80.0	272	10	US-09-903-456-80
8	36	80.0	272	10	US-09-903-456-81
9	36	80.0	272	12	US-10-156-911-75
10	36	80.0	272	12	US-10-156-911-76
11	36	80.0	272	12	US-10-156-911-78
12	36	80.0	272	12	US-10-156-911-80
13	36	80.0	272	12	US-10-156-911-81
14	36	80.0	283	10	US-09-903-456-29
15	36	80.0	283	10	US-09-903-456-88

16	36	80.0	283	12	US-10-156-911-29	Sequence 29, Appl
17	36	80.0	283	12	US-10-156-911-88	Sequence 88, Appl
18	36	80.0	283	12	US-10-408-736-26	Sequence 26, Appl
19	36	80.0	288	10	US-09-903-456-84	Sequence 84, Appl
20	36	80.0	288	12	US-10-156-911-84	Sequence 84, Appl
21	36	80.0	292	10	US-09-903-456-82	Sequence 82, Appl
22	36	80.0	292	12	US-10-156-911-82	Sequence 82, Appl
23	36	80.0	295	10	US-09-903-456-86	Sequence 86, Appl
24	36	80.0	295	12	US-10-156-911-86	Sequence 86, Appl
25	36	80.0	299	10	US-09-903-456-66	Sequence 66, Appl
26	36	80.0	299	12	US-10-156-911-66	Sequence 66, Appl
27	36	80.0	327	12	US-10-108-260A-3982	Sequence 3982, Ap
28	36	80.0	750	12	US-10-120-801-10	Sequence 10, Appl
29	36	80.0	851	12	US-10-108-260A-4227	Sequence 4227, Ap
30	35	77.8	76	12	US-09-864-408A-6578	Sequence 6578, Ap
31	35	77.8	119	12	US-09-864-408A-7798	Sequence 7798, Ap
32	35	77.8	140	12	US-10-094-749-3159	Sequence 3159, Ap
33	35	77.8	291	12	US-10-374-780A-302	Sequence 302, App
34	35	77.8	377	12	US-10-287-274-429	Sequence 429, App
35	35	77.8	420	12	US-10-369-493-3766	Sequence 3766, Ap
36	35	77.8	613	12	US-10-369-493-2509	Sequence 2509, Ap
37	35	77.8	667	12	US-10-369-493-23387	Sequence 23387, A
38	35	77.8	668	9	US-09-730-989-20	Sequence 20, Appl
39	35	77.8	1591	9	US-09-864-761-37952	Sequence 37952, A
40	34	75.6	62	9	US-09-864-761-35731	Sequence 35731, A
41	34	75.6	209	12	US-10-104-047-3213	Sequence 3213, Ap
42	34	75.6	216	12	US-09-833-245-1114	Sequence 1114, Ap
43	34	75.6	216	12	US-09-833-245-1115	Sequence 1115, Ap
44	34	75.6	236	12	US-09-976-782-106	Sequence 106, App
45	34	75.6	263	15	US-10-050-704-171	Sequence 171, App

ALIGNMENTS

RESULT 1
US-09-901-187B-10
; Sequence 10, Application US/09901187B
; Patent No. US20020151464A1
; GENERAL INFORMATION:
; APPLICANT: Panacea Pharmaceuticals, Inc.
; APPLICANT: Woloizin, Benjamin
; APPLICANT: Ostretova-Golts, Natalie
; APPLICANT: Lebowitz, Micheal S.
; TITLE OF INVENTION: Methods for Preventing Neural Tissue Damage and for the Treatment
; TITLE OF INVENTION: Alpha-Synuclein Diseases
; FILE REFERENCE: PAN01/002US
; CURRENT APPLICATION NUMBER: US/09/901,187B
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/217,319
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/279,199
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-901-187B-10

Query Match 100.0%; Score 45; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|||
Db 1 HLYHHKT 7

RESULT 2
US-09-903-456-68
; Sequence 68, Application US/09903456

```
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-68

Query Match      80.0%; Score 36; DB 10; Length 210;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      |:|||||
Db      167 HVYHHAT 173

RESULT 3
US-10-156-911-68
; Sequence 68, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-68

Query Match      80.0%; Score 36; DB 12; Length 210;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      |:|||||
Db      167 HVYHHAT 173

; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-75

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      |:|||||
Db      152 HVYHHAT 158

RESULT 4
US-09-903-456-75
; Sequence 75, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-75

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      |:|||||
Db      152 HVYHHAT 158

RESULT 5
US-09-903-456-76
; Sequence 76, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-76

Query Match      80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 HLYHHKT 7
      |:|||||
Db      152 HVYHHAT 158
```

RESULT 6
US-09-903-456-78
; Sequence 78, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-78

Query Match 80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHAT 158

RESULT 7
US-09-903-456-80
; Sequence 80, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-80

Query Match 80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHAT 158

RESULT 8
US-09-903-456-81
; Sequence 81, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-09-903-456-81

Query Match 80.0%; Score 36; DB 10; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHAT 158

RESULT 9
US-10-156-911-75
; Sequence 75, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-75

Query Match 80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

RESULT 10
US-10-156-911-76
; Sequence 76, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-76

Query Match 80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

RESULT 11
US-10-156-911-78
; Sequence 78, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-78

Query Match 80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

RESULT 12
US-10-156-911-80
; Sequence 80, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-80

Query Match 80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|||||
Db 152 HVYHHAT 158

RESULT 13
US-10-156-911-81
; Sequence 81, Application US/10156911
; Publication No. US20030163845A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407.US.P4
; CURRENT APPLICATION NUMBER: US/10/156,911
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 81
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
US-10-156-911-81

Query Match      80.0%; Score 36; DB 12; Length 272;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHHKT 7
      |:|:|:|
Db      152 HVYHHAT 158

RESULT 14
US-09-903-456-29
; Sequence 29, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Potential Mammalian Elongase
; NAME/KEY: VARIANT
; LOCATION: (282)...(282)
; OTHER INFORMATION: Xaa = Unknown or Other at position 282
US-09-903-456-29

Query Match      80.0%; Score 36; DB 10; Length 283;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHHKT 7
      |:|:|:|
Db      125 HVYHHAT 131

RESULT 15
US-09-903-456-88
; Sequence 88, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-Yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US.P3
; CURRENT APPLICATION NUMBER: US/09/903,456
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
```

```
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Thraustochytrium aureum
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (255)...(255)
; OTHER INFORMATION: Xaa = Unknown or other at position 255
US-09-903-456-88

Query Match      80.0%; Score 36; DB 10; Length 283;
Best Local Similarity 71.4%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1 HLYHHKT 7
      |:|:|:|
Db      134 HVYHHAT 140

Search completed: February 11, 2004, 17:54:13
Job time : 25.6667 secs
```

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:55:29 ; Search time 9.33333 Seconds
(without alignments)
72.127 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYVHKKT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	86.7	305	2 H82702	small conductance
2	37	82.2	212	2 B83866	hypothetical prote
3	36	80.0	91	2 H86701	hypothetical prote
4	36	80.0	114	2 S37150	asr2 protein - tom
5	36	80.0	115	2 T06588	abscisic stress ri
6	36	80.0	138	2 T02663	abscisic acid- and
7	36	80.0	169	2 T02081	ABA- and ripening-
8	36	80.0	259	2 T01976	hypothetical prote
9	36	80.0	325	2 T01975	hypothetical prote
10	35	77.8	242	2 G69633	glutamine transpor
11	35	77.8	292	2 T00829	wuschel protein -
12	35	77.8	367	2 D95196	oxidoreductase, Gf
13	35	77.8	367	2 A98063	conserved hypothet
14	35	77.8	367	2 AG0634	probable oxidoredu
15	35	77.8	376	2 AI2176	homocitrate syntha
16	35	77.8	377	2 S56505	hypothetical 41.9K
17	35	77.8	379	2 F95153	oxidoreductase, Gf
18	35	77.8	384	2 T49084	hypothetical prote
19	35	77.8	387	2 A82649	hypothetical prote
20	35	77.8	667	2 S66017	formate dehydrogen
21	35	77.8	982	2 T43699	DNA mismatch repai
22	35	77.8	3345	2 T13423	hypothetical prote
23	34	75.6	217	2 AI2186	hypothetical prote
24	34	75.6	365	2 T26226	hypothetical prote
25	34	75.6	383	2 F88486	protein F20H11.5 [
26	34	75.6	515	2 I60800	calcitonin recepto
27	34	75.6	515	2 I49154	calcitonin recepto
28	34	75.6	529	2 T00677	hypothetical prote
29	34	75.6	616	2 JQ1441	hypothetical 67K p

B. subtilis protei
B. subtilis protei
conserved hypothet
transcription regu
protein from bacte
probable MADS-box
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote

30 73.3 85 2 AC1180
31 73.3 85 2 AD1537
32 73.3 183 2 G89872
33 73.3 187 2 AH1346
34 73.3 236 2 E97101
35 73.3 256 2 B84688
36 73.3 285 1 C70066
37 73.3 287 2 S58648
38 73.3 295 2 T08183
39 73.3 304 2 A89905
40 73.3 309 2 T29293
41 73.3 310 2 T33497
42 73.3 320 2 T01874
43 73.3 392 2 T25583
44 73.3 392 2 T25209
45 73.3 425 2 T25457

ALIGNMENTS

RESULT 1

H82702
small conductance mechanosensitive ion channel XF1258 [imported] - Xylella fastidiosa (C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: H82702
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing, Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; PMID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82702
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <SIM>
A;Cross-references: GB:AE003960; GB:AE003849; NID:g9106242; PIDN:AAF84067.1; GSPDB:GN00
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y. , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santeili, R.V.; Sawasa A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1258

Query Match 86.7%; Score 39; DB 2; Length 305;
Best Local Similarity 85.7%; Pred. No. 8.3;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYVHKKT 7
|||
Db 282 HLYVHDT 288

RESULT 2

B83866
hypothetical protein BH1730 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83866
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83866
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <STO>
A;Cross-references: GB:AP001513; GB:BA000004; NID:g10174345; PIDN:BA05449.1; GSPDB:GN001513
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1730
C;Superfamily: Bacillus subtilis hypothetical protein ypjP

Query Match 82.2%; Score 37; DB 2; Length 212;
Best Local Similarity 85.7%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|||
Db 153 HLYHQKT 159

RESULT 3
H86701
hypothetical protein ygaE [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: H86701
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich, S.; Artamonov, A.; et al. 2001
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: H86701
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <STO>
A;Cross-references: GB:AE005176; PID:g12723513; PIDN:AAK04714.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ygaE

Query Match 80.0%; Score 36; DB 2; Length 91;
Best Local Similarity 85.7%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|||
Db 56 HLCCHKT 62

RESULT 4
S37150
asr2 protein - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C;Accession: S37150
R;Amitai, H.; Scolnik, P.A.; Bar-Zvi, D.
submitted to the EMBL Data Library, September 1993
A;Reference number: S37150
A;Accession: S37150
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <AMI>
A;Cross-references: EMBL:X74907; NID:g400468; PID:g400469
C;Genetics:
A;Introns: 53/3

Query Match 80.0%; Score 36; DB 2; Length 114;
Best Local Similarity 83.3%; Pred. No. 9.9;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 11 HLFHHK 16

RESULT 5
T06588
abscisic stress ripening protein 1 - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C;Accession: T06588
R;Iusem, N.D.; Bartholomew, D.M.; Hitz, W.D.; Scolnik, P.A.
Plant Physiol. 102, 1353-1354, 1993
A;Title: Tomato (Lycopersicon esculentum) transcript induced by water deficit and ripen
A;Reference number: Z15778; MUID:94105353; PMID:8278555
A;Accession: T06588
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-115 <IUS>
A;Cross-references: EMBL:L08255; NID:g170375; PIDN:AAA34137.1; PID:g170376
A;Experimental source: strain Ailsa Craig; fruit pericarp

Query Match 80.0%; Score 36; DB 2; Length 115;
Best Local Similarity 83.3%; Pred. No. 10;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 10 HLFHHK 15

RESULT 6
T02663
abscisic acid- and stress-induced protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: T02663
R;Vaidyanathan, R.; Kuruvilla, S.; Thomas, G.
submitted to the EMBL Data Library, December 1997
A;Description: Abscisic acid and stress inducible cDNA.
A;Reference number: Z14690
A;Accession: T02663
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-138 <VAI>
A;Cross-references: EMBL:AF039573; NID:g2773153; PID:g2773154
A;Experimental source: cultivar Pokkali, Vytilla 1
C;Genetics:
A;Gene: ASR1

Query Match 80.0%; Score 36; DB 2; Length 138;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
|||
Db 9 HLFHHK 14

RESULT 7
T02081
ABA- and ripening-inducible-like protein - maize
C;Species: Zea mays (maize)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 29-Oct-1999
C;Accession: T02081
R;Arredondo-Peter, R.; Shearman, L.; Ji, L.; Klucas, R.V.
submitted to the EMBL Data Library, April 1994
A;Description: Nucleotide sequence of an ABA- and ripening-like cDNA isolated from corn
A;Reference number: Z14553
A;Accession: T02081
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-169 <ARR>
A;Cross-references: EMBL:U09276; NID:g551482; PIDN:AAA21866.1; PID:g551483
A;Experimental source: strain Golden Bantam; mesophyll

Query Match 80.0%; Score 36; DB 2; Length 169;

Best Local Similarity 83.3%; Pred. No. 15;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHK 6
Db 10 HLFHHK 15
RESULT 8
T01976
hypothetical protein T9A4.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C;Accession: T01976
R;Zidanic, M.; McQuerry, Y.; Smith, A.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of A. thaliana T9A4.
A;Reference number: Z14478
A;Accession: T01976
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-259 <ZID>
A;Cross-references: EMBL:AF096373; NID:g3695400; PID:g3695410
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 48/3; 79/3; 142/3; 194/3
A;Note: T9A4.10
Query Match 80.0%; Score 36; DB 2; Length 259;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 74 HTFHHKT 80

RESULT 9
T01975
hypothetical protein T9A4.9 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Mar-1999
C;Accession: T01975
R;Zidanic, M.; McQuerry, Y.; Smith, A.
submitted to the EMBL Data Library, October 1998
A;Description: The sequence of A. thaliana T9A4.
A;Reference number: Z14478
A;Accession: T01975
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-325 <ZID>
A;Cross-references: EMBL:AF096373; NID:g3695400; PID:g3695413
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 62/3; 93/3; 156/3; 207/3
A;Note: T9A4.9
Query Match 80.0%; Score 36; DB 2; Length 325;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HLYHHKT 7
Db 88 HTFHHKT 94

RESULT 10
G69633
glutamine transport protein glnQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001

C;Accession: G69633
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: G69633
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-242 <KUN>
A;Cross-references: GB:Z99117; GB:Z99118; GB:AL009126; NID:g2635200; PID:CAB14701.1; F
A;Experimental source: strain 168
C;Genetics:
A;Gene: glnQ
C;Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology
C;Keywords: ATP; glutamine transport; nucleotide binding; P-loop
F;17-212/Domain: ATP-binding cassette homology <ABC>
F;34-41/Region: nucleotide-binding motif A (P-loop)
F;157-161/Region: nucleotide-binding motif B

Query Match 77.8%; Score 35; DB 2; Length 242;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 87 HLYPHKT 93

RESULT 11
T00829
wuschel protein - Arabidopsis thaliana
N;Alternate names: protein T13L16.3
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 30-Apr-1999
C;Accession: T00829
R;de la Bastide, M.; Hameed, A.; Gnoj, L.; Jensen, K.; Shohdy, N.; Gottesman, T.; Haber
McCombie, W.R.
submitted to the EMBL Data Library, January 1999
A;Description: A. thaliana BAC T13L16 from chromosome IV, top arm.
A;Reference number: Z14205
A;Accession: T00829
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-292
A;Cross-references: EMBL:AC003952; NID:g2708736; PID:g2708739
A;Experimental source: cultivar Columbia
C;Genetics:
A;Gene: wuschel
A;Map position: 2
A;Introns: 166/1; 195/3
A;Note: T13L16.3

Query Match 77.8%; Score 35; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
Db 152 HLYHH 156

RESULT 12

D95196
oxidoreductase, Gfo/Idh/MocA family SP1686 [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: D95196
R;Tetteilin, H.; Nelson, K.E.; Paulsen, I.T.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf, I.E.
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: D95196
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK75765.1; PID:gl4973179; GSPDB:GN00164; TIGR:SP4
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP1686

Query Match 77.8%; Score 35; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
|||
Db 170 HLYHH 174

RESULT 13

A98063
conserved hypothetical protein spr1530 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C;Accession: A98063
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: A98063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-367 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAL00334.1; PID:gl5459193; GSPDB:GN00174
C;Genetics:
A;Gene: spr1530

Query Match 77.8%; Score 35; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
|||
Db 170 HLYHH 174

RESULT 14

AG0634
probable oxidoreductase STY1170 [imported] - Salmonella enterica subsp. enterica serovar
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0634
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:43 ; Search time 5.16667 Seconds
(without alignments)
63.714 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHKT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	36	80.0	40	1 UC12_MAIZE	P80618 zea mays (m
2	36	80.0	114	1 ASR2_LYCES	P37219 lycopersico
3	36	80.0	115	1 ASR1_LYCES	Q08655 lycopersico
4	35	77.8	367	1 YG86_STRPN	Q54728 streptococc
5	35	77.8	372	1 YJHC_ECOLI	P39353 escherichia
6	35	77.8	376	1 NIV2_ANASP	P58637 anabaena sp
7	35	77.8	667	1 YYAE_BACSU	P37519 bacillus su
8	35	77.8	835	1 AXN1_BRARE	P57094 brachydanio
9	35	77.8	954	1 K6P1_YARLI	P59680 yarrowia li
10	35	77.8	982	1 MSH2_SCHPO	O74773 schizosacch
11	35	77.8	2483	1 PCX_DROME	P18490 drosophila
12	34	75.6	270	1 ELO3_HUMAN	Q9hb03 homo sapien
13	34	75.6	271	1 ELO3_MOUSE	Q35949 mus musculu
14	34	75.6	312	1 ELO4_MOUSE	Q9eqc4 mus musculu
15	34	75.6	314	1 ELO4_HUMAN	Q9gzr5 homo sapien
16	34	75.6	314	1 ELO4_MACFA	Q95k73 macaca fasc
17	34	75.6	326	1 RSP2_HUMAN	P50749 homo sapien
18	34	75.6	515	1 CALR_MOUSE	Q60755 mus musculu
19	34	75.6	516	1 CALR_RAT	P32214 rattus norv
20	33	73.3	259	1 MOB2_YEAST	P43563 saccharomyc
21	33	73.3	453	1 ZO6_XENLA	P18749 xenopus lae
22	33	73.3	802	1 ENAH_MOUSE	Q03173 mus musculu
23	33	73.3	841	1 AXN_CHICK	O42400 gallus gall
24	33	73.3	913	1 IF38_HUMAN	Q99613 homo sapien
25	33	73.3	2146	1 INSR_DROME	P09208 drosophila
26	32	71.1	292	1 ELO2_MOUSE	Q9jlj4 mus musculu
27	32	71.1	296	1 ELO2_HUMAN	Q9nxb9 homo sapien
28	32	71.1	312	1 NRT2_CHICK	P55807 gallus gall
29	32	71.1	326	1 YA55_METJA	Q58455 methanococc
30	32	71.1	331	1 Y011_NPVOP	Q65359 orgyia pseu
31	32	71.1	360	1 GBF2_ARATH	P42775 arabidopsis
32	32	71.1	399	1 EFTU_HELPJ	Q9zkl9 helicobacte
33	32	71.1	435	1 YYS3_CABEL	P49191 caenorhabdi

34	32	71.1	461	1 THIC_SYNEL	Q8dlz2 synechococc
35	32	71.1	518	1 TDT_MONDO	O02789 monodelphis
36	32	71.1	580	1 SYQ_RALSO	Q8y199 ralstonia s
37	32	71.1	581	1 Z319_MOUSE	Q9er18 mus musculu
38	32	71.1	582	1 Z319_HUMAN	Q9p2f9 homo sapien
39	32	71.1	595	1 THIC_BACHD	Q9kb14 bacillus ha
40	32	71.1	610	1 GLMS_PSEAE	Q9ht25 p glucosami
41	32	71.1	793	1 GAC1_YEAST	P28006 saccharomyc
42	32	71.1	842	1 AXN_XENLA	Q9ygy0 xenopus lae
43	32	71.1	937	1 PAC4_RAT	Q63415 rattus norv
44	32	71.1	969	1 PAC4_HUMAN	P29122 homo sapien
45	31	68.9	91	1 YVBE_VACCV	P20545 vaccinia vi

ALIGNMENTS

RESULT 1
UC12_MAIZE
ID UC12_MAIZE STANDARD; PRT; 40 AA.
AC P80618;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 237)
DE (Fragments).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.

TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
genome analysis program."
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.0, ITS MW IS: 21.7 kDa.
CC -!- SIMILARITY: TO THE TOMATO ABSICISIC STRESS RIPENING PROTEINS.
CC -!- CAUTION: THE ORDER OF THE PEPTIDES SHOWN IS UNCERTAIN.
DR Maize-2DPAGE; P80618; COLEOPTILE.
DR MaizeDB; 123942; -.
FT NON_TER 1
FT NON_CONS 15
FT NON_CONS 25
FT NON_TER 40
FT NON_TER 40
SQ SEQUENCE 40 AA; 4426 MW; 2E4B14B6B37B34C2 CRC64;
Query Match 80.0%; Score 36; DB 1; Length 40;
Best local Similarity 83.3%; Pred. No. 1.6;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHKK 6
|||:
Db 20 HLFHKK 25

RESULT 2
ASR2_LYCES
ID ASR2_LYCES STANDARD; PRT; 114 AA.
AC P37219;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Abscicisic stress ripening protein 2.
GN ASR2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.

```

OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RX MEDLINE=95148753; PubMed=7846175;
RA Amitai-Zeigerson H., Scolnik P.A., Bar-Zvi D.;
RT "Genomic nucleotide sequence of tomato Asr1 gene family.";
RL Plant Physiol. 106:1699-1700(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74907; CAAS2873.1; --
DR PIR; S37150; S37150.
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
FT DOMAIN 108 113 POLY-HIS.
FT DOMAIN 108 113 POLY-HIS.
SQ SEQUENCE 114 AA; 13020 MW; AE12FBBBCD3631248 CRC64;

Query Match      80.0%; Score 36; DB 1; Length 114;
Best Local Similarity 83.3%; Pred. No. 4.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
Db 11 HLFHHK 16

RESULT 3
ASR1_LYCES
ID ASR1_LYCES STANDARD; PRT; 115 AA.
AC Q08655;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Abscissic stress ripening protein 1.
GN ASR1.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RX MEDLINE=94105353; PubMed=8278555;
RA Iusem N.D., Bartholomew D.M., Hitz W.D., Scolnik P.A.;
RT "Tomato (Lycopersicon esculentum) transcript induced by water deficit
RT and ripening.";
RL Plant Physiol. 102:1353-1354(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Ailsa Craig;
RA Gilad A., Amitai-Zeigerson H., Scolnik P.A., Bar-Zvi D.;
RT "Genomic nucleotide sequence of the tomato stress/ripening induced
RT Asr1 gene.";
RL (In) Plant Gene Register PGR97-042.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: BY LEAF WATER DEFICIT AND FRUIT RIPENING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08255; AAA34137.1; --
DR EMBL; U86130; AAB64185.1; --
DR PIR; T06588; T06588.
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
KW Nuclear protein.
FT DOMAIN 6 10 POLY-HIS.
FT DOMAIN 74 78 POLY-ALA.
SQ SEQUENCE 115 AA; 13130 MW; 0575CC68A73FA176 CRC64;

Query Match      80.0%; Score 36; DB 1; Length 115;
Best Local Similarity 83.3%; Pred. No. 4.8;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
Db 10 HLFHHK 15

RESULT 4
YG86_STRPN
ID YG86_STRPN STANDARD; PRT; 367 AA.
AC Q54728;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase SP1686 (EC 1.1.1.1).
GN SP1686.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angioli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
RN [2]
RP SEQUENCE OF 1-347 FROM N.A.
RC STRAIN=Serotype 6;
RX MEDLINE=96326329; PubMed=8759848;
RA Berry A.M., Lock R.A., Paton J.C.;
RT "Cloning and characterization of nanB, a second Streptococcus
RT pneumoniae neuraminidase gene, and purification of the NanB enzyme
RT from recombinant Escherichia coli.";
RL J. Bacteriol. 178:4854-4860(1996).
CC -!- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY. STRONG, TO E.COLI
CC YJHC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE007461; AAK75765.1; --
DR EMBL; U43526; AAC44397.1; --
DR PIR; D95196; D95196.
DR TIGR; SP1686; --

```

DR InterPro; IPR000683; GFO_IDH_MoCA.
DR InterPro; IPR004104; GFO_IDH_MoCA_C.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
DR Pfam; PF02894; GFO_IDH_MoCA_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
FT CONFLICT 9 T -> A (IN REF. 2).
FT CONFLICT 156 P -> L (IN REF. 2).
SQ SEQUENCE 367 AA; 41095 MW; DE1D666A7752325D CRC64;

Query Match 77.8%; Score 35; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
DB 170 HLYHH 174

RESULT 5
YJHC_ECOLI STANDARD; PRT; 372 AA.
AC P39353; 1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical oxidoreductase yjhc (EC 1.1.1.1).
GN YJHC OR B4280.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes";
RL Nucleic Acids Res. 23:2105-2119(1995).
CC -!- SIMILARITY: BELONGS TO THE GFO/IDH/MoCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14003; AAA97176.1; ALT INIT.
DR EMBL; AE000498; AAC77236.1; ALT_INIT.
DR EcoGene; EG12545; yjhc.
DR InterPro; IPR000683; GFO_IDH_MoCA.
DR InterPro; IPR004104; GFO_IDH_MoCA_C.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
DR Pfam; PF02894; GFO_IDH_MoCA_C; 1.
KW Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 372 AA; 41384 MW; 34B2E535EADC2ABE CRC64;

Query Match 77.8%; Score 35; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
DB 170 HLYHH 174

RESULT 6
NIV2_ANASP STANDARD; PRT; 376 AA.
ID NIV2
AC P58637;

DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Homocitrate synthase 2 (EC 2.3.3.14).
GN NIPV2 OR ALR2968.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpō S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: THIS PROTEIN IS A FE-MO-COFACTOR BIOSYNTHETIC
CC COMPONENT.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + H(2)O + 2-oxoglutarate = 2-
CC hydroxybutane-1,2,4-tricarboxylate + CoA.
CC -!- SIMILARITY: Belongs to the alpha-IPM synthetase / homocitrate
CC synthase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003591; BAB74667.1; --
DR PIR; A12176; A12176.
DR InterPro; IPR002034; AIPM/Hcit_synth.
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF00682; HMGL-like; 1.
DR PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
DR PROSITE; PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
KW Nitrogen fixation; Transferase; Complete proteome.
SQ SEQUENCE 376 AA; 40336 MW; 343A804D990E4300 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
DB 241 HLYHH 245

RESULT 7
YYAE_BACSU STANDARD; PRT; 667 AA.
ID YYAE_BACSU
AC P37519;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yyae.
GN YYAE.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin";
RL DNA Res. 1:1-14(1994).


```
ID AC K6P1_YARLI STANDARD; PRT; 954 AA.
AD P59680;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE 6-phosphofructokinase alpha subunit (EC 2.7.1.11) (Phosphofructokinase
DE 1) (Phosphohexokinase) (6PF-1-K alpha subunit).
GN PFK1.
OS Yarrowia lipolytica (Candida lipolytica);
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RA Flores C.L., Martinez-Costa O.H., Sanchez V., Aragon J.J.,
RA Cancedo C.;
RT "The phosphofructokinase from the yeast Yarrowia lipolytica.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
CC fructose 1,6-bisphosphate.
CC -!- ENZYME REGULATION: Allosterically inhibited by ATP and activated
CC by AMP and fructose 2,6-bisphosphate (By similarity).
CC -!- PATHWAY: Key control step of glycolysis.
CC -!- SUBUNIT: Heterooctamer of 4 alpha and 4 beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the phosphofructokinase family. Two domains
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY142710; AAN34943.1;
DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 2.
KW Kinase; Transferase; Glycolysis; Repeat; Allosteric enzyme.
SQ SEQUENCE 954 AA; 104049 MW; 33CD38C292F2450 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 954;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHT 7
DB 896 HLYAHT 902

RESULT 10
MSH2 SCHPO STANDARD; PRT; 982 AA.
ID O74773; O42950;
AC O74773; O42950;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein msh2.
GN MSH2 OR SPBC19G7.01C OR SPBC24C6.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99077962; PubMed=9858548;
RA Rudolph C., Kunz C., Parisi S., Lehmann E., Hartsuiker E.,
RA Fartmann B., Kramer W., Kohli J., Fleck O.;
RT "The msh2 gene of Schizosaccharomyces pombe is involved in mismatch
RT repair, mating-type switching, and meiotic chromosome organization.";
RL Mol. Cell. Biol. 19:241-250(1999).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckie E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Involved in post-replicative DNA-mismatch repair. Binds
CC to mismatch-containing DNA.
CC -!- FUNCTION: Required for correct termination of copy synthesis
CC during mating-type switching. Also required for proper chromosome
CC organization during meiosis.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ006948; CAA07342.1;
DR EMBL; AL031786; CAA21156.1;
DR EMBL; AL021839; CAA17055.1;
DR PIR; T43699; T43699.
DR GenDB SPombe; SPBC19G7.01c;
DR InterPro; IPR000432; Muts_C.
DR InterPro; IPR002863; Muts_N.
DR Pfam; PF01624; Muts_I; 1.
DR Pfam; PF05188; Muts_II; 1.
DR Pfam; PF05192; Muts_III; 1.
DR Pfam; PF05190; Muts_IV; 1.
DR Pfam; PF00488; Muts_V; 1.
DR ProDom; PD001263; Muts_C; 1.
DR SMART; SM00534; Mutsac; 1.
DR SMART; SM00533; Mutsd; 1.
DR PROSITE; PS00486; DNA_MISMATCH_REPAIR_2; 1.
KW DNA repair; ATP-binding; DNA-binding; Nuclear protein.
FT NP_BIND 721 728 ATP (POTENTIAL).
SQ SEQUENCE 982 AA; 109738 MW; 6A5F5AF711620244 CRC64;

Query Match 77.8%; Score 35; DB 1; Length 982;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHH 5
DB 304 HLYHH 308
```

QY 1 HLYHHK 6
Db 356 HLHHK 361

RESULT 12
ELO3_HUMAN
ID ELO3_HUMAN STANDARD; PRT; 270 AA.
AC Q9HB03; Q8N180;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Elongation of very long chain fatty acids protein 3 (Cold inducible
DE Glycoprotein of 30 kDa).
GN ELOVL3 OR CIG30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE OF 35-270 FROM N.A.
RA Semina E.V., Murray J.C.;
RT "The CIG30/PITX3/GBF1 gene configuration is conserved between human
RT and mouse genomes."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: May be involved in a membrane event related to cellular
CC proliferation in brown adipose tissue. Could be implicated in
CC synthesis of very long chain fatty acids and sphingolipids. May
CC catalyze one or both of the reduction reaction in fatty acid
CC elongation, i.e., conversion of beta-ketoacyl CoA to beta-
CC hydroxyacyl CoA or reduction of trans-2-enoyl CoA to the saturated
CC acyl CoA derivative (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -1- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC034344; AAH34344.1; -.
CC EMBL; AF292387; AAG17875.1; -.
CC Genew; HGNC:18047; ELOVL3.
CC InterPro; IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.

RESULT 11
PCX_DROME
ID PCX_DROME STANDARD; PRT; 2483 AA.
AC P18490;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Pecanex protein.
GN PCX.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE OF 1-545 FROM N.A.
RC STRAIN=Oregon-R;
RA Labonne S.G.;
RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 546-2483 FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=90033754; PubMed=2478400;
RA Labonne S.G., Sunitha I., Mahowald A.P.;
RT "Molecular Genetics of pecanex, a maternal-effect neurogenic locus of
RT Drosophila melanogaster that potentially encodes a large
RT transmembrane protein."
RL Dev. Biol. 136:1-16 (1989).
CC -1- FUNCTION: INVOLVED IN NEUROGENESIS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74329; AAA28747.1; -.
CC EMBL; M25662; AAA28749.1; -.
CC FIR; A37361; A37361.
CC FlyBase; FBgn0003048; pcx.
CC Pfam; PF05041; Pecanex_C; 1.
CC Developmental protein; Neurogenesis; Transmembrane; Glycoprotein;
KW Repeat.
FT TRANSMEM 750 771 POTENTIAL.
FT TRANSMEM 859 883 POTENTIAL.
FT TRANSMEM 1285 1307 POTENTIAL.
FT DOMAIN 39 51 POLY-HIS.
FT DOMAIN 1231 1240 5 X 2 AA TANDEM REPEATS OF G-T.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 711 711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1118 1118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1337 1337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1350 1350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1528 1528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1688 1688 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1786 1786 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2315 2315 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 2483 AA; 266639 MW; D376A3AB57112AE CRC64;

Query Match 77.8%; Score 35; DB 1; Length 2483;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DR PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSNMEM 29 49 POTENTIAL.
FT TRANSNMEM 66 86 POTENTIAL.
FT TRANSNMEM 164 186 POTENTIAL.
FT TRANSNMEM 198 218 POTENTIAL.
FT TRANSNMEM 235 255 POTENTIAL.
FT SITE 266 268 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF (POTENTIAL).
SQ SEQUENCE 270 AA; 31500 MW; 0C6C8F1E7B5DE8B1 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 270;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 145 HWYHST 151

RESULT 13
ELO3_MOUSE STANDARD; PRT; 271 AA.
AC O35949;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Elongation of very long chain fatty acids protein 3 (Cold inducible glycoprotein of 30 kDa) (CIN-2).
GN ELOVL3 OR CIG30.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brown adipose tissue;
RX MEDLINE=98058971; PubMed=9395518;
RA Tvrdik P., Asadi A., Kozak L.P., Nedergaard J., Cannon B., Jacobsson A.;
RT "Cig30, a mouse member of a novel membrane protein gene family, is involved in the recruitment of brown adipose tissue."
RL J. Biol. Chem. 272:31738-31746(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=99403085; PubMed=10473596;
RA Tvrdik P., Asadi A., Kozak L.P., Nuglozeh E., Parente F., Nedergaard J., Jacobsson A.;
RT "Cig30 and Pitx3 genes are arranged in a partially overlapping tail-to-tail array resulting in complementary transcripts."
RL J. Biol. Chem. 274:26387-26392(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT PROSITE; PS01188; ELO; 1.
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -|- FUNCTION: May be involved in a membrane event related to cellular proliferation in brown adipose tissue. Could be implicated in synthesis of very long chain fatty acids and sphingolipids. May catalyze one or both of the reduction reaction in fatty acid elongation, i.e., conversion of beta-ketoacyl CoA to beta-hydroxyacyl CoA or reduction of trans-2-enoyl CoA to the saturated acyl CoA derivative.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum (Potential).
CC -|- TISSUE SPECIFICITY: Readily detected in brown adipose tissue and liver. Weak expression in kidney, white adipose tissue, heart and skin. Not detected in lung, testis, muscle, spleen, brain, thymus and intestine.
CC -|- INDUCTION: Elevated in brown adipose tissue in conditions of brown fat recruitment, namely cold stress, perinatal development and after diet-induced thermogenesis. A synergistic action of both catecholamines and glucocorticoids is required for the induction.
CC -|- SIMILARITY: BELONGS TO THE ELO FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U97107; AAC06127.1; --
DR EMBL; AF054504; AAD51088.1; --
DR EMBL; BC016468; AAH16468.1; --
DR MGD; MGI:1195976; Elov13.
DR InterPro; IPR002076; GNS1_SUR4.
DR Pfam; PF01151; ELO; 1.
DR PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum;
KW Glycoprotein.
FT TRANSNMEM 30 50 POTENTIAL.
FT TRANSNMEM 67 87 POTENTIAL.
FT TRANSNMEM 165 187 POTENTIAL.
FT TRANSNMEM 199 219 POTENTIAL.
FT TRANSNMEM 236 256 POTENTIAL.
FT SITE 269 269 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 271 AA; 32060 MW; F7CA96199BE89401 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 271;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 146 HWYHST 152

RESULT 14
ELO4_MOUSE STANDARD; PRT; 312 AA.
AC Q9EQC4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation of very long chain fatty acids protein 4. ELOVL4.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=20578755; PubMed=11138005;
RA Zhang K., Kniazeva M., Han M., Li W., Yu Z., Yang Z., Li Y.,
RA Metzker M.L., Allikmets R., Zack D.J., Kakuk L.E., Lagali P.S.,
RA Wong P.W., McDonald I.M., Sieving P.A., Figueroa D.J., Austin C.P.,
RA Gould R.J., Ayyagari R., Petrukhin K.;
RT "A 5-bp deletion in ELOVL4 is associated with two related forms of
RT autosomal dominant macular dystrophy.";
RL Nat. Genet. 27:89-93(2001).
CC -!- FUNCTION: Involved in the biosynthesis of very long chain fatty
CC acids. Seems to represent a photoreceptor-specific component of
CC the fatty acid elongation system residing on the endoplasmic
CC reticulum.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in the adult retina, exclusively in
CC photoreceptor cells.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF277093; AAG47667.1; -.
CC MGD; MGI:1933331; Elov14.
CC InterPro; IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum.
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT SITE 308 310
FT ENDOPASMIC RETICULUM RETRIEVAL MOTIF
FT (POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 312 AA; 36520 MW; B04CD48024772132 CRC64;

Query Match 75.6%; Score 34; DB 1; Length 312;
Best Local Similarity 71.4%; Pred. No. 30;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
|:|:|:|
Db 158 HVYHCT 164

RESULT 15
ELO4 HUMAN
ID ELO4 HUMAN STANDARD; PRT; 314 AA.
AC Q9GZR5; Q9H139;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation of very long chain fatty acids protein 4.
GN ELOVL4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT VAL-299.
RC TISSUE=Retina;
RX MEDLINE=20578755; PubMed=11138005;
RA Zhang K., Kniazeva M., Han M., Li W., Yu Z., Yang Z., Li Y.,
RA Metzker M.L., Allikmets R., Zack D.J., Kakuk L.E., Lagali P.S.,
RA Wong P.W., McDonald I.M., Sieving P.A., Figueroa D.J., Austin C.P.,
RA Gould R.J., Ayyagari R., Petrukhin K.;

RT "A 5-bp deletion in ELOVL4 is associated with two related forms of
RT autosomal dominant macular dystrophy.";
RL Nat. Genet. 27:89-93(2001).
RX MEDLINE=21464738; PubMed=11581213;
RA Edwards A.O., Donoso L.A., Ritter R. III;
RA "A novel gene for autosomal dominant Stargardt-like macular dystrophy
RT with homology to the SUR4 protein family.";
RT Invest. Ophthalmol. Vis. Sci. 42:2652-2663(2001).
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 35-314 FROM N.A.
RA Dunn M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in the biosynthesis of very long chain fatty
CC acids. Seems to represent a photoreceptor-specific component of
CC the fatty acid elongation system residing on the endoplasmic
CC reticulum. May be implicated in docosahexaenoic acid (DHA)
CC biosynthesis, which requires dietary consumption of the essential
CC alpha-linolenic acid and a subsequent series of three elongation
CC steps. May be involved in one of these three elongation steps.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in the retina and at much lower
CC level in the brain.
CC -!- DISEASE: Defects in ELOVL4 are the cause of Stargardt disease 3
CC macular dystrophy (STGD3) and autosomal dominant macular dystrophy
CC (ADMD). Both diseases are autosomal dominant inherited forms of
CC macular degeneration characterized by decreased visual acuity,
CC macular atrophy, and extensive fundus flecks.
CC -!- SIMILARITY: BELONGS TO THE ELO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF279654; AAG47669.1; -.
CC EMBL; AF279649; AAG47669.1; JOINED.
CC EMBL; AF279650; AAG47669.1; JOINED.
CC EMBL; AF279651; AAG47669.1; JOINED.
CC EMBL; AF279652; AAG47669.1; JOINED.
CC EMBL; AF279653; AAG47669.1; JOINED.
CC EMBL; AF277094; AAG47668.1; -.
CC EMBL; AY037298; AAK68639.1; -.
CC EMBL; AK055277; BAB70895.1; -.
CC EMBL; AL132875; CAC19496.1; -.
CC Genew; HGNC:14415; ELOVL4.
CC MIM; 605512; -.
CC MIM; 600110; -.
CC GO; GO:0008020; F:G-protein coupled photoreceptor activity; NAS.
CC GO; GO:0006633; P:fatty acid biosynthesis; NAS.
CC InterPro; IPR002076; GNS1_SUR4.
CC Pfam; PF01151; ELO; 1.
CC PROSITE; PS01188; ELO; 1.
KW Fatty acid biosynthesis; Transmembrane; Endoplasmic reticulum;
KW Polymorphism; Stargardt disease; Vision.

FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 247 267 POTENTIAL.
FT SITE 310 312 ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
(POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 299 299 M -> V.
FT SQ SEQUENCE 314 AA; 36829 MW; B2EBCE54D868E96E CRC64;
/FTid=VAR_012492.

Query Match 75.6%; Score 34; DB 1; Length 314;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
|:|:|:|
Db 158 HVYHCT 164

Search completed: February 11, 2004, 17:04:14
Job time : 6.16667 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:52:34 ; Search time 24.5833 Seconds
(without alignments)
73.479 Million cell updates/sec

Title: US-09-901-187C-10
Perfect score: 45
Sequence: 1 HLYHHKT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	86.7	305	16 Q9PDX1	Q9pdx1 xylella fas
2	38	84.4	191	10 Q9FM60	Q9fm60 arabidopsis
3	37	82.2	212	16 Q9KC44	Q9kc44 bacillus ha
4	37	82.2	221	5 Q9VTF0	Q9vtf0 drosophila
5	37	82.2	222	5 Q8IQF0	Q8iqf0 drosophila
6	36	80.0	91	16 Q9CHV2	Q9chv2 lactococcus
7	36	80.0	103	10 Q40165	Q40165 lycopersico
8	36	80.0	108	10 Q9ZRB5	Q9zrb5 solanum tub
9	36	80.0	109	10 Q9ZRB6	Q9zrb6 solanum tub
10	36	80.0	110	10 Q82575	Q82575 lycopersico
11	36	80.0	138	10 Q49149	Q49149 oryza sativ
12	36	80.0	142	10 Q948L3	Q948l3 saccharum o
13	36	80.0	149	10 Q94G23	Q94g23 vitis vinif
14	36	80.0	169	10 Q41730	Q41730 zea mays (m
15	36	80.0	218	11 Q8BHM5	Q8bhm5 mus musculu
16	36	80.0	259	10 Q82620	Q82620 arabidopsis

17	36	80.0	278	10 Q9SQV0	Q9sqv0 arabidopsis
18	36	80.0	282	13 Q8AX86	Q8ax86 brachydanio
19	36	80.0	298	10 Q9SQU9	Q9squ9 arabidopsis
20	36	80.0	299	11 Q920L7	Q920l7 rattus norv
21	36	80.0	299	11 Q8R5D3	Q8r5d3 mus musculu
22	36	80.0	299	11 Q8BUE3	Q8bue3 mus musculu
23	36	80.0	299	11 Q8BLG6	Q8blg6 mus musculu
24	36	80.0	299	11 Q8BHI7	Q8bhi7 mus musculu
25	36	80.0	325	10 Q82619	Q82619 arabidopsis
26	36	80.0	327	4 Q8N8G4	Q8n8g4 homo sapien
27	36	80.0	350	5 Q9NE06	Q9ne06 leishmania
28	36	80.0	540	4 Q8IYP3	Q8iyp3 homo sapien
29	36	80.0	851	4 Q8N861	Q8n861 homo sapien
30	36	80.0	4832	5 Q8I1Q7	Q8i1q7 plasmodium
31	35	77.8	140	4 Q96M18	Q96m18 homo sapien
32	35	77.8	242	16 Q34677	Q34677 bacillus su
33	35	77.8	277	5 Q9VCZ0	Q9vcz0 drosophila
34	35	77.8	291	10 Q9SB92	Q9sb92 arabidopsis
35	35	77.8	292	10 Q9SL51	Q9sl51 arabidopsis
36	35	77.8	293	13 Q8AWE8	Q8awe8 oreochromis
37	35	77.8	367	2 Q9F924	Q9f924 pasteurella
38	35	77.8	367	16 Q9CK74	Q9ck74 pasteurella
39	35	77.8	367	16 Q8Z7M7	Q8z7m7 salmonella
40	35	77.8	367	16 Q8DNU5	Q8dnu5 streptococc
41	35	77.8	368	16 Q8ZQ31	Q8zq31 salmonella
42	35	77.8	379	2 Q9L4J9	Q9l4j9 anabaena va
43	35	77.8	379	16 Q97QA0	Q97qa0 streptococc
44	35	77.8	384	10 Q9SUZ9	Q9suz9 arabidopsis
45	35	77.8	387	16 Q9PC54	Q9pcs4 xylella fas

ALIGNMENTS

RESULT 1
Q9PDX1 PRELIMINARY; PRT; 305 AA.
ID
AC Q9PDX1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Small conductance mechanosensitive ion channel.
GN XF1258.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159(2000).
DR EMBL; AE003960; AAF84067.1; --
DR InterPro; IPR006685; MSion channel.
DR Pfam; PF00924; MS_channel; 1.
KW Complete proteome.
SQ SEQUENCE 305 AA; 33648 MW; 10F62B51496A2D53 CRC64;

Query Match 86.7%; Score 39; DB 16; Length 305;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 282 HLYHHKT 288

RESULT 2
Q9FM60 PRELIMINARY; PRT; 191 AA.
AC Q9FM60;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Genomic DNA, chromosome 5, p1 clone:MDF20.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RT Sequence features of the regions of 1,456,315 bp covered by nineteen
RT physically assigned p1 and TAC clones.";
RL DNA Res. 5:41-54(1998).
DR EMBL; AB09050; BAB09246.1; --
SQ SEQUENCE 191 AA; 21713 MW; DEDF9AD52940E056 CRC64;

Query Match 84.4%; Score 38; DB 10; Length 191;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 50 HLYHHGT 56

RESULT 3
Q9KC44 PRELIMINARY; PRT; 212 AA.
AC Q9KC44;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BH1730.
GN BH1730.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; AP001513; BAB05449.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 212 AA; 24141 MW; E0A75F501BDD9B2 CRC64;

Query Match 82.2%; Score 37; DB 16; Length 212;
Best Local Similarity 85.7%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHKT 7
Db 153 HLYHQT 159

RESULT 4
Q9VTF0 PRELIMINARY; PRT; 221 AA.
AC Q9VTF0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG11801 protein.
GN CG11801.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Paclele J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003546; AAF50101.2; -;
DR FlyBase; FBgn0036128; CG11801.
DR InterPro; IPR002076; GNS1_SUR4.
DR Pfam; PF01151; ELO; 1.
DR PROSITE; PS01188; ELO; 1.
SQ SEQUENCE 221 AA; 26345 MW; F0DF5FA5A299F365 CRC64;

Query Match 82.2%; Score 37; DB 5; Length 221;
Best Local Similarity 71.4%; Pred. NO. 34;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
|:|:|:|
Db 99 HVYHST 105

RESULT 5
Q8IQF0 PRELIMINARY; PRT; 222 AA.
ID Q8IQF0;
AC Q8IQF0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG32072-PA.
GN CG32072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bröttier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Paclele J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003546; AAF50101.2; -;
DR FlyBase; FBgn0036128; CG11801.
DR InterPro; IPR002076; GNS1_SUR4.
DR Pfam; PF01151; ELO; 1.
DR PROSITE; PS01188; ELO; 1.
SQ SEQUENCE 221 AA; 26345 MW; F0DF5FA5A299F365 CRC64;

Query Match 82.2%; Score 37; DB 5; Length 221;
Best Local Similarity 71.4%; Pred. NO. 34;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
|:|:|:|
Db 99 HVYHST 105

RESULT 5
Q8IQF0 PRELIMINARY; PRT; 222 AA.
ID Q8IQF0;
AC Q8IQF0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG32072-PA.
GN CG32072.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
FT NON_TER 103
SQ SEQUENCE 103 AA; 11631 MW; 926ED85B07660C6F CRC64;

Query Match 80.0%; Score 36; DB 10; Length 103;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
Db 11 HLFHHK 16

RESULT 8
Q9ZRB5 PRELIMINARY; PRT; 108 AA.
AC Q9ZRB5;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Ci21B protein.
GN Ci21B.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Saturna;
RA Schneider A., Salamini F., Gebhardt C.;
RT "Expression patterns and promoter activity of the cold regulated gene ci21A of potato.";
RL Plant Physiol. 0:0-0(1996).
DR EMBL; U76611; AAD00255.1; -;
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.
SQ SEQUENCE 108 AA; 12159 MW; 03CE2ECF8179BA30 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 108;
Best Local Similarity 83.3%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
Db 9 HLFHHK 14

RESULT 9
Q9ZRB6 PRELIMINARY; PRT; 109 AA.
AC Q9ZRB6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Ci21A protein.
GN Ci21A.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Saturna;
RA Schneider A., Salamini F., Gebhardt C.;
RT "Expression patterns and promoter activity of the cold regulated gene ci21A of potato.";
RL Plant Physiol. 0:0-0(1996).
DR EMBL; U76610; AAD00254.1; -;
DR InterPro; IPR003496; ABA_WDS.
DR Pfam; PF02496; ABA_WDS; 1.

Query Match 82.2%; Score 37; DB 5; Length 222;
Best Local Similarity 71.4%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 7
Db 100 HLYHHK 106

RESULT 6
Q9CHV2 PRELIMINARY; PRT; 91 AA.
AC Q9CHV2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Unknown protein.
GN YGAE OR IL0616.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006294; AAK04714.1; -;
KW Complete proteome.
SQ SEQUENCE 91 AA; 10517 MW; A299E9C1682C5BFD CRC64;

Query Match 80.0%; Score 36; DB 16; Length 91;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHHK 7
Db 56 HLYHHK 62

RESULT 7
Q40165 PRELIMINARY; PRT; 103 AA.
AC Q40165;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE ABA- and ripening-induced protein (Fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94218393; PubMed=8165244;
RA Rossi M., Iusem N.D.;
RT "Tomato (Lycopersicon esculentum) genomic clone homologous to a gene encoding an abscisic acid-induced protein.";
RL Plant Physiol. 104:1073-1074(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93352822; PubMed=7626782;
RA Rossi M., Iusem N.D.;
RT "Sequence of Asr2, a member of a gene family from Lycopersicon esculentum encoding chromosomal proteins: homology to an intron of the polygalacturonase gene.";
RL DNA Seq. 5:225-227(1995).
DR EMBL; L20756; AAA99440.1; -.

SQ SEQUENCE 109 AA; 12429 MW; 0981E6F4509F0135 CRC64;

Query Match 80.0%; Score 36; DB 10; Length 109;

Best Local Similarity 83.3%; Pred. No. 26;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

||:||||

Db 9 HLFHHK 14

RESULT 10

O82575 PRELIMINARY; PRT; 110 AA.

AC O82575;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Fruit-ripening protein.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asteridae; lamids; Solanales; Solanaceae; Solanum.

OX NCBI_TaxID=4081;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AVRDC CL5915-93D4-1-0-3;

RA Wang Y.-C., Hsieh H.-L.;

RT "Nucleotide sequence of a cDNA encoding a tomato fruit-ripening

protein.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF093141; AAC61780.1; -.

DR InterPro; IPR003496; ABA_WDS.

DR Pfam; PF02496; ABA_WDS; 1.

SQ SEQUENCE 110 AA; 12555 MW; 467416108F74F363 CRC64;

Query Match

Best Local Similarity 80.0%; Score 36; DB 10; Length 110;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

||:||||

Db 10 HLFHHK 15

RESULT 11

O49149

AC O49149 PRELIMINARY; PRT; 138 AA.

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Abscissic acid-and stress-inducible protein.

GN ASR1.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzeae; Oryza.

OX NCBI_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Pokkali;

RA Vaidyanathan R., Kuruvilla S., Thomas G.;

RT "Characterization and expression pattern of an abscissic acid and

osmotic stress responsive gene from rice.";

RL Plant Sci. 140:25-36(1999).

DR EMBL; AF039573; AAB96681.1; -.

DR Gramene; O49149; -.

DR InterPro; IPR003496; ABA_WDS.

DR Pfam; PF02496; ABA_WDS; 1.

SQ SEQUENCE 138 AA; 15465 MW; 3EE0C8E99C641D0E CRC64;

Query Match

80.0%; Score 36; DB 10; Length 138;

Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

||:||||

Db 9 HLFHHK 14

RESULT 12

Q948L3

ID Q948L3 PRELIMINARY; PRT; 142 AA.

AC Q948L3;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Drought inducible 22 kDa protein.

GN SODIP22.

OS Saccharum officinarum (Sugarcane).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC PACAD clade; Panicoideae; Andropogoneae; Saccharum.

OX NCBI_TaxID=4547;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=M442-51;

RA Sugiharto B., Ermawati N., Mori H., Aoki K., Sakakibara K.Y.,

RA Yamaya T., Sugiyama T., Sakakibara H.;

RT "Identification and characterization of a gene encoding drought-

inducible protein that localizes in the bundle sheath cell of

sugarcane.";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB071694; BAB68268.1; -.

DR InterPro; IPR003496; ABA_WDS.

DR Pfam; PF02496; ABA_WDS; 1.

SQ SEQUENCE 142 AA; 15932 MW; 35DC5BC66D6A75A1 CRC64;

Query Match

80.0%; Score 36; DB 10; Length 142;

Best Local Similarity 83.3%; Pred. No. 34;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6

||:||||

Db 9 HLFHHK 14

RESULT 13

Q94G23

ID Q94G23 PRELIMINARY; PRT; 149 AA.

AC Q94G23;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Putative transcription factor.

OS Vitis vinifera (Grape).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;

OC Vitis.

OX NCBI_TaxID=29760;

RN [1]

RP SEQUENCE FROM N.A.

RA Atanassova R.A., Cakir B., Gaillard C., Delrot S.;

RT "A transcription factor binding to the promoter of a grape berry

hexose transporter.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF281656; AAK69513.1; -.

DR InterPro; IPR003496; ABA_WDS.

DR Pfam; PF02496; ABA_WDS; 1.

SQ SEQUENCE 149 AA; 16703 MW; D46C6B82953F66E5 CRC64;

Query Match

80.0%; Score 36; DB 10; Length 149;

Best Local Similarity 83.3%; Pred. No. 35;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Job time : 27.5833 secs

QY 1 HLYHHK 6
||:||||
Db 9 HLFHHK 14

RESULT 14
Q41730
ID Q41730 PRELIMINARY; PRT; 169 AA.
AC Q41730;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
DE ABA- and ripening-inducible-like protein.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Arredondo-Peter R., Shearman L., Ji L., Klucas R.V.;
RT "Nucleotide sequence of an ABA- and ripening-like cDNA isolated from
RT corn roots.";
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U09276; AAA21866.1; --
SQ SEQUENCE 169 AA; 18503 MW; 024AF2AA3D07B71F CRC64;

Query Match 80.0%; Score 36; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HLYHHK 6
||:||||
Db 10 HLFHHK 15

RESULT 15
Q8BHM5
ID Q8BHM5 PRELIMINARY; PRT; 218 AA.
AC Q8BHM5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
DE Hypothetical FYVE Zn-finger.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK049341; BAC33694.1; --
DR EMBL; AK086031; BAC39597.1; --
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 24303 MW; 03306ADFF63E85BB CRC64;

Query Match 80.0%; Score 36; DB 11; Length 218;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HLYHKT 7
||:||||
Db 186 HLYHKT 192

Search completed: February 11, 2004, 17:09:44

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:51:33 ; Search time 32.25 Seconds
(without alignments)
34.452 Million cell updates/sec

Title: US-09-901-187C-11
Perfect score: 44
Sequence: 1 THIHPS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	7	AAE14556	Human alpha-synuclein
2	43	97.7	148	ABP57617	S. murayamaensis A
3	39	88.6	178	AAU19257	Human G protein-co
4	39	88.6	248	ABG18105	Novel human diagno
5	39	88.6	301	ABG19187	Novel human diagno
6	39	88.6	301	ABG21623	Novel human diagno
7	39	88.6	322	ABG21650	Novel human diagno
8	39	88.6	334	ABG21635	Novel human diagno
9	39	88.6	383	ABG05684	Novel human diagno

10	39	88.6	383	22	ABG19176	Novel human diagno
11	39	88.6	402	14	AAE14556	Transcriptase, Ba
12	39	88.6	437	22	ABG02252	Novel human diagno
13	39	88.6	451	22	ABG19205	Novel human diagno
14	39	88.6	455	22	ABG10309	Novel human diagno
15	39	88.6	478	22	ABG21601	Novel human diagno
16	39	88.6	495	22	ABG05178	Novel human diagno
17	39	88.6	495	22	ABG19192	Novel human diagno
18	39	88.6	505	22	ABG21655	Novel human diagno
19	39	88.6	517	22	AAU31031	Novel human secret
20	39	88.6	524	22	ABG02251	Novel human diagno
21	39	88.6	524	22	ABG19196	Novel human diagno
22	39	88.6	524	22	ABG25240	Novel human diagno
23	39	88.6	524	22	ABG29174	Novel human diagno
24	39	88.6	536	22	ABG02250	Novel human diagno
25	39	88.6	546	22	ABG19203	Novel human diagno
26	39	88.6	548	22	ABG21453	Novel human diagno
27	39	88.6	568	22	ABG28701	Novel human diagno
28	39	88.6	601	22	ABG02256	Novel human diagno
29	39	88.6	632	22	ABG17784	Novel human diagno
30	39	88.6	645	22	ABG02255	Novel human diagno
31	39	88.6	646	23	AAU11293	Ashbya gossypii AG
32	39	88.6	665	22	ABG04325	Novel human diagno
33	39	88.6	673	22	ABG02254	Novel human diagno
34	39	88.6	728	22	ABG19197	Novel human diagno
35	39	88.6	786	22	ABG02257	Novel human diagno
36	39	88.6	839	22	ABG22934	Novel human diagno
37	39	88.6	863	22	ABG19208	Novel human diagno
38	39	88.6	1067	22	ABE12481	Human bone marrow
39	39	88.6	1130	22	ABG02249	Novel human diagno
40	39	88.6	1455	22	ABG02259	Novel human diagno
41	39	88.6	1481	22	ABG13697	Novel human diagno
42	38	86.4	779	22	AAE95562	Human protein sequ
43	38	86.4	1377	22	ABE64706	Drosophila melanog
44	38	86.4	1878	19	AAW81170	Human BAZ2-alpha p
45	38	86.4	1878	22	AAE40239	Human polypeptide

ALIGNMENTS

RESULT 1

AAE14556
ID AAE14556 standard; peptide; 7 AA.

XX AAE14556;

AC AAE14556;

DT 17-MAY-2002 (first entry)

XX Human alpha-synuclein aggregation inhibitor #11.
DE Alpha-synuclein; inhibitor; neurodegenerative disease; Lewy body;
KW Parkinson's disease; Alzheimer's disease; diffuse Lewy body disease;
KW multiple system atrophy; Hallervorden-Spatz disease; human.

XX Homo sapiens.

XX WO200204482-A1.

PN 17-JAN-2002.

PD 06-JUL-2001; 2001WO-US21379.

PR 07-JUL-2000; 2000US-217319P.

PR 28-MAR-2001; 2001US-279199P.

PA (PANA-) PANACEA PHARM INC.

PI Wolozin B, Ostretova-Golts N, Lebowitz MS;

XX WPI; 2002-179695/23.

XX Determination of an agent capable of inhibiting aggregation of alpha

PT synuclein useful for treating a neurodegenerative disease involves
PT determining aggregation of alpha synuclein in the presence of exogenous
PT iron or copper
XX
PS Claim 40; Page 37; 52pp; English.
XX
CC The invention relates to screening of inhibitors of alpha-synuclein
CC aggregation in the presence of exogenous iron or copper. The inhibitors
CC are magnesium and alpha-synuclein binding peptides, which are
CC useful for treating neurodegenerative disease that involves
CC the formation of Lewy bodies e.g. Parkinson's disease (PD), Alzheimer's
CC disease (AD), diffuse Lewy body disease, mixed AD-PD, multiple system
CC atrophy and Hallervorden-Spatz disease. The present sequence is a
CC peptide that binds to the NAC (non-amyloid-beta protein component)
CC portion of human alpha-synuclein and inhibits its aggregation.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 44; DB 23; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.3e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHHP 7
Db 1 THIHHP 7

RESULT 2
ID ABP57617 standard; Protein; 148 AA.

AC ABP57617;
XX
DT 29-APR-2003 (first entry)
XX
DE S. murayamaensis ATCC 21414 kinamycin protein SEQ ID NO:19.
XX

KW Streptomyces murayamaensis ATCC 21414; Streptomyces murayamaensis;
KW glycosylated kinamycin; kinamycin; type II polyketide; polyketide;
KW antibacterial; cytostatic; infection; antibiotic; antitumour;
KW electrophilic azo-coupling agent.

OS Streptomyces murayamaensis.

XX WO2003002066-A2.

XX 09-JAN-2003.

XX 27-JUN-2002; 2002WO-US20719.

XX 27-JUN-2001; 2001US-301401P.

XX (DIVE-) DIVERSA CORP.

PI Short JM, Paradkar A, Varoglu M, Mathur EJ;

XX WPI; 2003-210195/20.
DR N-PSDB; ABZ71140.

XX New isolated polyketide used e.g. as antibiotic and antitumor agents
PT comprises kinamycin molecule comprising at least one saccharide group

XX Claim 77; Page 92; 119pp; English.

XX The present invention describes a polyketide comprising a kinamycin
CC molecule comprising at least one saccharide group. ABZ71132 to ABZ71163
CC encode glycosylated kinamycins ABP57609 to ABP57640 isolated from
CC Streptomyces murayamaensis ATCC 21414. Kinamycins are a class of
CC type II polyketides. The kinamycins have antibacterial and cytostatic
CC activities. They can be used for treating infections as antibiotics
CC and as antitumour agents, and as electrophilic azo-coupling agents
CC in vitro or in vivo.

XX SQ Sequence 148 AA;
Query Match 97.7%; Score 43; DB 24; Length 148;
Best Local Similarity 85.7%; Pred. No. 2.4;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHHP 7
Db 93 THVHHP 99

RESULT 3
AAU19257
ID AAU19257 standard; Protein; 178 AA.

XX AAU19257;

DT 04-DEC-2001 (first entry)

DE Human G protein-coupled receptor nGPCR-2434.

XX
KW Human; G protein-coupled receptor; nGPCR-x; antiviral; analgesic;
KW cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive;
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;
KW schizophrenia; migraine; major depression; anxiety; mental disorder;
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.

XX Homo sapiens.

XX WO200166750-A2.

XX 13-SEP-2001.

XX 08-MAR-2001; 2001WO-US07322.

XX 08-MAR-2000; 2000US-0187581.

PR 08-MAR-2000; 2000US-0187582.

PR 08-MAR-2000; 2000US-0187714.

PR 08-MAR-2000; 2000US-0187715.

PR 08-MAR-2000; 2000US-0187825.

PR 08-MAR-2000; 2000US-0187828.

PR 08-MAR-2000; 2000US-0187829.

PR 08-MAR-2000; 2000US-0187830.

PR 08-MAR-2000; 2000US-0187833.

PR 08-MAR-2000; 2000US-0187874.

PR 08-MAR-2000; 2000US-0187930.

PR 08-MAR-2000; 2000US-0188049.

PR 08-MAR-2000; 2000US-0189294.

PR 08-MAR-2000; 2000US-0187929.

PR 08-MAR-2000; 2000US-0187928.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS;

XX WPI; 2001-536778/59.

XX N-PSDB; AAS30826.

XX Isolated nucleic acid molecules encoding G protein-coupled receptors
PT termed nGPCR-x, useful in the treatment and diagnosis of viral
PT infections, cancers and mental disorders (e.g. Parkinson's disease and
PT schizophrenia) -
XX Claim 31; Page 312-313; 336pp; English.

XX The invention relates to novel isolated nucleic acid molecules encoding
CC G protein-coupled receptors termed nGPCR-x. nGPCR-x polynucleotides,
CC polypeptides, and modulators may be used in the treatment of diseases and

CC conditions such as infections, such as viral infections caused by HIV-1
CC (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and
CC cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity,
CC anorexia, hypotension, hypertension, myocardial infarction,
CC atherosclerosis), Parkinson's disease, and psychotic and
CC neurological disorders, including schizophrenia, migraine, major
CC depression, anxiety, mental disorder, manic depression, and
CC dyskinesias, such as Huntington's disease or Tourette's Syndrome
CC and many other diseases and syndromes listed in the specification.
CC nGPCR-x polynucleotides and polypeptides, as well as nGPCR-x
CC modulators, may also be used in diagnostic assays for such diseases or
CC conditions. The present sequence represents a G protein-coupled
CC receptor of the invention.

XX Sequence 178 AA;
SQ Query Match 88.6%; Score 39; DB 22; Length 178;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHPS 7
Db 102 HIHHPS 107

RESULT 4
ABG18105
ID ABG18105 standard; Protein; 248 AA.

XX AC ABG18105;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #18096.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS82292.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX PS Claim 20; SEQ ID No 48464; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 248 AA;

Query Match 88.6%; Score 39; DB 22; Length 248;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHHPS 7
Db 61 HIHHPS 66

RESULT 5
ABG19187
ID ABG19187 standard; Protein; 301 AA.

XX AC ABG19187;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #19178.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS83374.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX PS Claim 20; SEQ ID No 49546; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 301 AA;
SQ Query Match 88.6%; Score 39; DB 22; Length 301;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 7
|||
Db 95 THCHPS 101

RESULT 6
ABG21623
ID ABG21623 standard; Protein; 301 AA.
XX
AC ABG21623;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21614.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85810.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 51982; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 301 AA;
SQ Query Match 88.6%; Score 39; DB 22; Length 301;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 THIHPS 7
|||
Db 95 THCHPS 101

RESULT 7
ABG21650
ID ABG21650 standard; Protein; 322 AA.
XX
AC ABG21650;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21641.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85837.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 52009; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 322 AA;
Query Match 88.6%; Score 39; DB 22; Length 322;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 95 THCHPS 101
RESULT 8
ABG21635
ID ABG21635 standard; Protein; 334 AA.
XX
AC ABG21635;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #21626.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85822.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 51994; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 334 AA;
Query Match 88.6%; Score 39; DB 22; Length 334;
Best Local Similarity 85.7%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 95 THCHPS 101
RESULT 9
ABG05684
ID ABG05684 standard; Protein; 383 AA.
XX
AC ABG05684;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5675.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS69871.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 36043; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 383 AA;

Query Match 88.6%; Score 39; DB 22; Length 383;

Best Local Similarity 85.7%; Pred. No. 37;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 10 THCHPS 16

RESULT 10

ABG19176

ID ABG19176 standard; Protein; 383 AA.

XX AC ABG19176;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #19167.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS83363.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity -

XX PS Claim 20; SEQ ID No 49535; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX CC to restore normal activity of (II) or to treat disease states involving

XX CC quantitating a polypeptide in tissue, as molecular weight markers and as

XX CC a food supplement. (II) and its binding partners are useful in medical

XX CC imaging of sites expressing (II). (I) and (II) are useful for treating

XX CC disorders involving aberrant protein expression or biological activity.

XX CC The polypeptide and polynucleotide sequences have applications in

XX CC diagnostics, forensics, gene mapping, identification of mutations

XX CC responsible for genetic disorders or other traits to assess biodiversity

XX CC and to produce other types of data and products dependent on DNA and

XX CC amino acid sequences. ABG00010-ABG30377 represent novel human

XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 383 AA;

Query Match 88.6%; Score 39; DB 22; Length 383;

Best Local Similarity 85.7%; Pred. No. 37;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 247 THCHPS 253

RESULT 11

AAR41310

ID AAR41310 standard; Protein; 402 AA.

XX AC AAR41310;

XX DT 26-APR-1994 (first entry)

XX DE Transcriptase.

XX KW Retrotransposon; transcriptase; gene destruction; ss.

XX OS Bacillus stearothermophilus.

XX PN JP05227974-A.

XX PD 07-SEP-1993.

XX PF 24-FEB-1992; 92JP-0036426.

XX PR 24-FEB-1992; 92JP-0036426.

XX PA (MARU-) MARUKAN SU KK.

XX DR WPI; 1993-316610/40.

XX DR N-PSDB; AAQ49161.

XX PT Retrotransposon contg.; gene encoding transcriptase - has base

XX PT sequence of C at position 10 to G at position 1339

XX PS Claim 2; Page 4-6; 7pp; Japanese.

XX CC The transcriptase gene, encoded by bases 118-1323 is claimed

XX CC (claim 4) as is the retrotransposon containing bases 10-1339 of

XX CC the sequence shown (claim 1). The transcriptase is useful for studying

XX CC the mechanism of gene destruction or expression in Bacillus species.

XX SQ Sequence 402 AA;

Query Match 88.6%; Score 39; DB 14; Length 402;

Best Local Similarity 85.7%; Pred. No. 39;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 247 THCHPS 253

RESULT 12

ABG02252

ID ABG02252 standard; Protein; 437 AA.

XX AC ABG02252;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #2243.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX WO200175067-A2.
PN
XX
PD
XX
XX 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS66439.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 32611; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 437 AA;
Query Match 88.6%; Score 39; DB 22; Length 437;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 247 THCHPS 253
RESULT 13
ABG19205
ID ABG19205 standard; Protein; 451 AA.
XX
AC ABG19205;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19196.
XX
KW Human; chromosome mapping; gene mapping; Gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.

XX 11-OCT-2001.
PN
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS83392.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 49564; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 451 AA;
Query Match 88.6%; Score 39; DB 22; Length 451;
Best Local Similarity 85.7%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHPS 7
Db 247 THCHPS 253
RESULT 14
ABG10309
ID ABG10309 standard; Protein; 455 AA.
XX
AC ABG10309;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #10300.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS74496.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 40668; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 455 AA;
Query Match 88.6%; Score 39; DB 22; Length 455;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHHP 7
Db 238 THCHHP 244
RESULT 15
ABG21601
ID ABG21601 standard; Protein; 478 AA.
XX AC ABG21601;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #21592.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS85788.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity -
XX PS Claim 20; SEQ ID No 51960; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human
XX CC diagnostic amino acid sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 478 AA;
Query Match 88.6%; Score 39; DB 22; Length 478;
Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 THIHHP 7
Db 323 THCHHP 329
Search completed: February 11, 2004, 17:03:04
Job time : 33.25 secs

OM protein - protein search, using sw model

Run on: February 11, 2004, 16:56:09 ; Search time 10.4167 Seconds
(without alignments)
28.433 Million cell updates/sec

Title: US-09-901-187C-11

Perfect score: 44

Sequence: 1 THIHPS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	39	88.6	646	4	US-09-625-188-10
2	35	79.5	253	4	US-09-252-991A-20052
3	35	79.5	315	4	US-09-252-991A-19219
4	35	79.5	441	4	US-09-252-991A-32115
5	35	79.5	613	4	US-09-328-352-7191
6	34	77.3	172	4	US-09-252-991A-22814
7	34	77.3	338	3	US-09-330-611-6
8	34	77.3	344	2	US-08-755-728-3
9	34	77.3	344	2	US-08-974-655-3
10	34	77.3	344	3	US-09-283-011-3
11	34	77.3	347	2	US-09-016-000-1
12	34	77.3	511	4	US-09-252-991A-28223
13	34	77.3	558	4	US-09-252-991A-21305
14	34	77.3	568	4	US-09-252-991A-27039
15	33	75.0	212	4	US-09-328-352-7678
16	33	75.0	255	4	US-09-031-962D-4
17	33	75.0	279	3	US-08-963-901-4
18	33	75.0	296	4	US-09-252-991A-18620
19	33	75.0	299	4	US-09-252-991A-17615
20	33	75.0	315	4	US-09-252-991A-20012
21	33	75.0	329	4	US-09-252-991A-18007
22	33	75.0	361	4	US-09-252-991A-27989
23	33	75.0	391	4	US-09-134-001C-5234
24	33	75.0	393	4	US-09-252-991A-19043
25	33	75.0	410	4	US-09-252-991A-25451
26	33	75.0	457	3	US-08-963-901-2
27	33	75.0	520	4	US-09-252-991A-24248

28	33	75.0	672	4	US-09-252-991A-24905	Sequence 24905, A
29	33	75.0	859	4	US-09-369-364A-5	Sequence 5, Appli
30	32	72.7	149	4	US-09-252-991A-19603	Sequence 19603, A
31	32	72.7	263	4	US-09-328-352-4326	Sequence 4326, Ap
32	32	72.7	449	2	US-08-927-394-2	Sequence 2, Appli
33	32	72.7	1315	4	US-09-252-991A-22746	Sequence 22746, A
34	32	72.7	3665	2	US-08-222-617A-13	Sequence 13, Appli
35	32	72.7	3712	2	US-08-222-617A-4	Sequence 4, Appli
36	32	72.7	3712	2	US-08-222-617A-25	Sequence 25, Appli
37	31	70.5	94	4	US-09-252-991A-31551	Sequence 31551, A
38	31	70.5	141	4	US-09-353-719-2	Sequence 2, Appli
39	31	70.5	217	4	US-09-252-991A-24723	Sequence 24723, A
40	31	70.5	292	4	US-08-858-207A-391	Sequence 391, App
41	31	70.5	357	4	US-09-252-991A-22404	Sequence 22404, A
42	31	70.5	456	4	US-09-328-352-4944	Sequence 4944, Ap
43	31	70.5	538	4	US-09-252-991A-20413	Sequence 20413, A
44	31	70.5	565	4	US-09-506-286B-8	Sequence 8, Appli
45	31	70.5	565	4	US-09-506-286B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-625-188-10
; Sequence 10, Application US/09625188
; Patent No. 6307037
; GENERAL INFORMATION:
; APPLICANT: No. 6307037artis AG
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-31285PI
; CURRENT APPLICATION NUMBER: US/09/625.188
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 646
; TYPE: PRT
; ORGANISM: Ashbya gossypii
US-09-625-188-10

Query Match 88.6%; Score 39; DB 4; Length 646;
Best Local Similarity 100.0%; Pred.No.31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 HIHPS 7

Db 304 HIHPS 309

RESULT 2

US-09-252-991A-20052
; Sequence 20052, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20052
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20052

Query Match 79.5%; Score 35; DB 4; Length 253;

Best Local Similarity 100.0%; Pred. No. 56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|:|||||
Db 26 HVHHP 30

RESULT 3
US-09-252-991A-19219
; Sequence 19219, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19219
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19219

Query Match 79.5%; Score 35; DB 4; Length 315;
Best Local Similarity 66.7%; Pred. No. 70;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 7
|:|||||
Db 307 HVHHPA 312

RESULT 4
US-09-252-991A-32115
; Sequence 32115, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32115
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (31)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-32115

Query Match 79.5%; Score 35; DB 4; Length 441;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 7
|:|||||
Db 277 HVHHPA 282

RESULT 5
US-09-328-352-7191
; Sequence 7191, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7191
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7191

Query Match 79.5%; Score 35; DB 4; Length 613;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 THIHHP 6
|:|||||
Db 579 THAHHP 584

RESULT 6
US-09-252-991A-22814
; Sequence 22814, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22814
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22814

Query Match 77.3%; Score 34; DB 4; Length 172;
Best Local Similarity 80.0%; Pred. No. 56;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHHP 6
|:|||||
Db 2 HVHHP 6

RESULT 7
US-09-330-611-6
; Sequence 6, Application US/09330611
; Patent No. 6248874
; GENERAL INFORMATION:
; APPLICANT: FREY, Perry A.
; APPLICANT: RUZICKA, Frank J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING BACTERIAL LYSINE 2,3-AMINOMUTASE
; FILE REFERENCE: 032026/0476
; CURRENT APPLICATION NUMBER: US/09/330,611
; CURRENT FILING DATE: 1999-06-11
; EARLIER APPLICATION NUMBER: US 09/198,942
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 338
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-330-611-6

Query Match 77.3%; Score 34; DB 3; Length 338;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 THIHPS 7

Db 225 THINHPN 231

RESULT 8

US-08-755-728-3
Sequence 3, Application US/08755728
Patent No. 5962312

GENERAL INFORMATION:

APPLICANT: Plowman, Gregory
APPLICANT: Mossie, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/755,728
FILING DATE: No. 5962312ember 25, 1996

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/008,809
FILING DATE: December 18, 1995
APPLICATION NUMBER: 60/023,943
FILING DATE: August 14, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 223/113
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-755-728-3

Query Match 77.3%; Score 34; DB 2; Length 344;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPS 7

Db 131 HLHPN 136

RESULT 9

US-08-974-655-3

Sequence 3, Application US/08974655

Patent No. 5972676

GENERAL INFORMATION:

APPLICANT: Plowman, Gregory
APPLICANT: Mossie, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,655

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/755,728
FILING DATE: No. 5972676ember 25, 1996
APPLICATION NUMBER: 60/008,809
FILING DATE: December 18, 1995
APPLICATION NUMBER: 60/023,943
FILING DATE: August 14, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 223/113
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO

US-08-974-655-3

Query Match 77.3%; Score 34; DB 2; Length 344;

Best Local Similarity 66.7%; Pred. No. 1.1e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHPS 7

Db 131 HLHPN 136

RESULT 10

US-09-283-011-3

Sequence 3, Application US/09283011

Patent No. 6207401

GENERAL INFORMATION:

APPLICANT: Plowman, Gregory

APPLICANT: Mossie, Kevin
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF AUR-1
TITLE OF INVENTION: AND/OR AUR-2 RELATED DISORDERS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/283,011
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/012,135
FILING DATE: January 22, 1998
APPLICATION NUMBER: 08/755,728
FILING DATE: No. 6207401ember 25, 1996
APPLICATION NUMBER: 60/023,943
FILING DATE: August 14, 1996
APPLICATION NUMBER: 60/008,809
FILING DATE: December 18, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 231/282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 344 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-283-011-3

Query Match 77.3%; Score 34; DB 3; Length 344;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7
|:|:|:
Db 131 HLHPN 136

RESULT 11
US-09-016-000-1
; Sequence 1, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,000
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0465 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMC1NOT01
CLONE: 2940
US-09-016-000-1

Query Match 77.3%; Score 34; DB 2; Length 347;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HIHPS 7
|:|:|:
Db 131 HLHPN 136

RESULT 12
US-09-252-991A-28223
; Sequence 28223, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28223
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28223

Query Match 77.3%; Score 34; DB 4; Length 511;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|:|:|
Db 318 HVHHP 322

RESULT 13
US-09-252-991A-21305
; Sequence 21305, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21305
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21305

Query Match 77.3%; Score 34; DB 4; Length 558;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 7
|:|:|
Db 345 HLHHP 350

RESULT 14
US-09-252-991A-27039
; Sequence 27039, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27039
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27039

Query Match 77.3%; Score 34; DB 4; Length 568;
Best Local Similarity 80.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|:|:|
Db 439 HVHHP 443

RESULT 15
US-09-328-352-7678
; Sequence 7678, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7678
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7678

Query Match 75.0%; Score 33; DB 4; Length 212;
Best Local Similarity 80.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 HHHHP 6
|:|:|
Db 100 HLHHP 104

Search completed: February 11, 2004, 17:13:39
Job time : 10.4167 secs